

## ИССЛЕДОВАНИЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ СЕТЕВЫХ СРЕЗОВ

Кузнецов А.А., Тарасов В.Н.

Поволжский государственный университет телекоммуникаций и информатики, Самара, РФ

E-mail: a1997@internet.ru, v.tarasov@psuti.ru

Данная статья посвящена исследованию алгоритмов машинного обучения для распознавания сетевых срезов в 5G. В условиях динамичного развития сетевых технологий сегментирование сети становится ключевым инструментом для оптимизации сетей. Автор анализирует различные подходы и методы машинного обучения, направленные на эффективное определение и классификацию сетевых срезов, что имеет критическое значение для управления сетевыми ресурсами и повышения общей производительности. Цель исследования заключается в анализе моделей для поиска такой из них, которая будет способна точно определять наиболее подходящие фрагменты сети для различных видов услуг, что, в свою очередь, способствует повышению производительности сети и эффективности использования ресурсов. Результаты и выводы исследования могут быть полезны для развития более эффективных и автоматизированных методов управления сетью на основе анализа сетевых срезов с использованием машинного обучения.

**Ключевые слова:** сетевые срезы, сетевая нарезка, технологии 5G, машинное обучение, сегментация сети

### Введение

Современные телекоммуникационные сети стремительно развиваются, и с появлением технологий пятого поколения (5G) возможности скоростной передачи данных и время отклика вышли на качественно новый уровень. Технологии 5G внедрили понятие сетевого сегментирования или сетевой нарезки (Network Slicing), которая позволяет на основе одной физической инфраструктуры создавать множество виртуализированных сетевых срезов, оптимизированных для обслуживания различных типов приложений и устройств [1]. Сетевая нарезка позволяет динамически адаптировать сеть под уникальные потребности и характеристики каждого типа трафика, обеспечивая высокое качество обслуживания для каждого пользователя. Однако создание, поддержание и управление этими сетевыми срезами является сложной задачей, требующей автоматизации на всех уровнях.

В сетях 5G технология сетевой нарезки стала не просто возможностью, а необходимостью для поддержки широкого спектра приложений, от расширенной мобильной широкополосной связи (eMBB) до ультра-надежной связи с низкими задержками (Ultra Reliable and Low Latency Communications, URLLC) и массовой межмашинной связи (mMTC). Каждый из этих типов приложений предъявляет особые требования к сети, такие как скорость передачи данных, задержка и надежность соединения, что делает технологию

нарезки сети важным аспектом оптимизации сетевых ресурсов и управления качеством обслуживания (QoS). Без эффективного управления сетевой нарезкой невозможно обеспечить высокую производительность и доступность сети 5G, что ставит под угрозу выполнение критических задач, таких как, например, работа умных городов, беспилотных автомобилей, требующих высоконадежных соединений и минимальных задержек. Сетевая нарезка позволяет провайдерам настраивать сеть в зависимости от запросов клиентов и их потребностей, разделяя физическую инфраструктуру на независимые виртуальные сети, каждая из которых адаптирована под специфический трафик или услугу. Например, одному пользователю может потребоваться высокоскоростное соединение для работы с мультимедиа, тогда как другому – стабильное и надежное соединение для IoT-устройств, передающих данные в реальном времени. Однако реализация и поддержка такой инфраструктуры связаны с рядом проблем [2]. Наиболее важная из них – автоматизация управления сетевыми срезами, что невозможно без применения наилучшего алгоритма машинного обучения.

### Проблемы и сложности сетевого сегментирования

Управление сегментированием сети включает в себя несколько задач, таких как выбор оптимального сетевого среза для различных типов трафика, распределение ресурсов, мониторинг производительности и адаптация к изменени-

ям в нагрузке. Сложность этой задачи повышается с увеличением количества подключенных устройств, объема данных и разнообразия требований к QoS [3]. В условиях высокой нагрузки и постоянно меняющихся требований к сети, традиционные методы управления уже не справляются с задачей оптимизации. Это приводит к необходимости внедрения интеллектуальных систем, которые могут анализировать большие объемы данных в реальном времени и принимать решения на основе полученных данных. Однако сетевая нарезка требует значительных вычислительных ресурсов и быстрой обработки данных для эффективного управления. Это делает традиционные алгоритмы неэффективными, поскольку они не могут адаптироваться к быстро меняющимся условиям сети и предоставлять высокоточные прогнозы и управление сетевыми срезами. Применение машинного обучения для сегментирования сети становится более актуальным, поскольку оно позволяет автоматизировать процесс анализа и управления, обеспечивая динамическое распределение ресурсов и адаптацию к изменениям в режиме реального времени [4]. Основная проблема состоит в необходимости выбора оптимального алгоритма для классификации и управления сетевыми срезами. На сегодняшний день существует множество алгоритмов машинного обучения, которые могут эффективно выполнять задачу сегментации, однако каждый из них имеет свои сильные и слабые стороны. Три из наиболее популярных методов – это метод опорных векторов (Support Vector Machine, SVM), случайные леса (Random Forest, RF) и деревья решений. Одной из наиболее сложных задач в исследовании является качественная оценка их работы в условиях приближенных к реальной работе сети. Поэтому важно правильно подготовить данные и корректно оценить показатели методов машинного обучения.

### Основные метрики качества для оценки алгоритмов

Каждый из методов имеет свои сильные стороны, и они по-разному справляются с одной и той же задачей в зависимости от условий. Определить, какой метод работает лучше, непросто из-за различных метрик, таких как точность, устойчивость к переобучению, скорость обработки и возможность адаптации к меняющимся данным [5]. Для объективной оценки требуется комплексный подход, который включает сравнение методов по ключевым метрикам. Точность и F1 Score – эти метрики показывают качество клас-

сификации, но они могут сильно варьироваться в зависимости от сложности данных и структуры модели [6].

Скорость обработки, которая в условиях сетевого сегментирования важна для управления скоростью работы алгоритма, чтобы обрабатывать данные в реальном времени. Этот фактор особенно важен для RF и SVM, так как все это может быть ресурсоемко. И возможность адаптации. Так как в сетях 5G данные изменяются в реальном времени, и алгоритмы должны адаптироваться к новым условиям [7]. Здесь требуется оценка устойчивости модели к изменению данных, что может быть сложным критерием для SVM.

### Подготовка данных для обучения

В этом исследовании был проанализирован набор данных, содержащий 31 583 строки и 17 столбцов, с уклоном на категории данных LTE/5G, IoT. Первичное исследование не выявило проблемных значений в наборе данных. Сводная статистика демонстрирует нормальное распределение набора данных, в диапазоне значений от 0 до 1 для различных категорий. Парные диаграммы, гистограммы, были использованы для исследования взаимосвязей, распределений и потенциальных артефактов. Анализ выявил сильные корреляции между характеристиками «IoT», «смартфон» и конкретными категориями смарт-технологии. Это исследование служит предварительным шагом к пониманию характеристик набора данных и может стать основой для дальнейшего углубленного анализа или моделирования. Каждая строка представляет собой конкретную запись данных, в столбцах содержится информация о времени, скорости потери пакетов, задержки пакетов и другие показатели. Значения в каждой строке отражают конкретные характеристики. Эти показатели отражают различные аспекты производительности и использования сети в различных секторах или приложениях, как представлено в таблице 1.

Таблица 1. Набор данных в используемом датасете

| Номер | Столбец          | Количество данных |
|-------|------------------|-------------------|
| 0     | LTE/5g Category  | 31583             |
| 1     | Time             | 31583             |
| 2     | Packet Loss Rate | 31583             |
| 3     | Packet delay     | 31583             |
| 4     | IoT              | 31583             |
| 5     | LTE/5G           | 31583             |
| 6     | GBR              | 31583             |
| 7     | Non-GBR          | 31583             |

(Продолжение) Таблица 1. Набор данных в используемом датасете

| Номер | Столбец              | Количество данных |
|-------|----------------------|-------------------|
| 8     | AR/VR/Gaming         | 31583             |
| 9     | Healthcare           | 31583             |
| 10    | Industry 4.0         | 31583             |
| 11    | IoT Devices          | 31583             |
| 12    | Public Safety        | 31583             |
| 13    | Smart City & Home    | 31583             |
| 14    | Smart Transportation | 31583             |
| 15    | Smartphone           | 31583             |
| 16    | slice Type           | 31583             |

В таблице 2 показаны статистические сводки для различных атрибутов в наборе данных. Он включает в себя подсчеты для каждого атрибута, показывая 31 583 записи.

Средние значения обозначают значение для категории LTE/5G (10,97), времени (11,48), скорости потери пакетов (0,003), задержки пакетов (114,13), остальные показатели указывают на средние пропорции или значения бинарных показателей для различных областей, таких как IoT и т. д. Показатель стандартного отклонения измеряет отклонение от среднего значения для дисперсии данных. Кроме того, в таблицу включены квартили (25-й, 50-й и 75-й), дающие представление о распределении и диапазоне данных, как показано в таблице 3.

Все это позволяет получить общее представление о распределении и диапазоне значений для различных типов сетевых срезов при различных показателях производительности и приложениях. Разобрав данные показатели с графиков, можно приступить к выбору алгоритма машинного обучения.

Таблица 2. Набор данных в используемом датасете. Детализация по каждой категории данных

| № набора данных      | 1     | 2        | 3     | 4     | 5     |
|----------------------|-------|----------|-------|-------|-------|
| LTE/5g Category      | 15    | 14       | 11    | 20    | 2     |
| Time                 | 17    | 18       | 7     | 14    | 22    |
| Packet Loss Rate     | 0,001 | 0,000001 | 0,001 | 0,001 | 0,001 |
| Packet delay         | 100   | 10       | 50    | 50    | 50    |
| IoT                  | 0     | 1        | 0     | 0     | 0     |
| LTE/5G               | 1     | 0        | 0     | 0     | 1     |
| GBR                  | 0     | 1        | 1     | 0     | 0     |
| Non-GBR              | 0     | 0        | 0     | 0     | 0     |
| AR/VR/Gaming         | 1     | 0        | 0     | 0     | 0     |
| Healthcare           | 1     | 1        | 0     | 0     | 0     |
| Industry 4.0         | 1     | 1        | 0     | 1     | 1     |
| IoT Devices          | 1     | 1        | 0     | 0     | 0     |
| Public Safety        | 0     | 1        | 0     | 0     | 1     |
| Smart City & Home    | 1     | 1        | 0     | 1     | 1     |
| Smart Transportation | 1     | 1        | 1     | 0     | 0     |
| Smartphone           | 1     | 0        | 0     | 0     | 1     |

Таблица 3. Набор данных в используемом датасете. Общие показатели

| Параметр             | count | mean       | std        | min      | 25%      | 50%   | 75%  | max  |
|----------------------|-------|------------|------------|----------|----------|-------|------|------|
| LTE/5g Category      | 31585 | 10,974923  | 6,053641   | 1        | 6        | 11    | 16   | 22   |
| Time                 | 31585 | 11,476459  | 6,915643   | 0        | 6        | 11    | 17   | 23   |
| Packet Loss Rate     | 31585 | 0,003079   | 0,004344   | 0,000001 | 0,000001 | 0,001 | 0,01 | 0,01 |
| Packet delay         | 31585 | 114,126745 | 106,320976 | 10       | 50       | 75    | 150  | 300  |
| IoT                  | 31585 | 0,4681     | 0,498989   | 0        | 0        | 0     | 1    | 1    |
| LTE/5G               | 31585 | 0,5319     | 0,498989   | 0        | 0        | 1     | 1    | 1    |
| GBR                  | 31585 | 0,441187   | 0,496537   | 0        | 0        | 0     | 1    | 1    |
| Non-GBR              | 31585 | 0,558813   | 0,496537   | 0        | 0        | 1     | 1    | 1    |
| AR/VR/Gaming         | 31585 | 0,106291   | 0,308215   | 0        | 0        | 0     | 0    | 1    |
| Healthcare           | 31585 | 0,057943   | 0,233639   | 0        | 0        | 0     | 0    | 1    |
| Industry 4.0         | 31585 | 0,116107   | 0,320358   | 0        | 0        | 0     | 0    | 1    |
| IoT Devices          | 31585 | 0,057879   | 0,233519   | 0        | 0        | 0     | 0    | 1    |
| Public Safety        | 31585 | 0,059716   | 0,236963   | 0        | 0        | 0     | 0    | 1    |
| Smart City & Home    | 31585 | 0,117722   | 0,322283   | 0        | 0        | 0     | 0    | 1    |
| Smart Transportation | 31585 | 0,058734   | 0          | 0        | 0        | 0     | 0    | 1    |
| Smartphone           | 31585 | 0,494443   | 0          | 0        | 0        | 0     | 1    | 1    |

## Описание процесса построения и оценки базовых моделей

Для того чтобы подготовить собранный набор исходных данных к эффективному использованию, требуется дополнительная нормализация данных. После предварительной обработки набор данных разбивается на три категории: данные для обучения, проверочные данные и данные для валидации результатов, с процентным соотношением 80:10:10. Это позволило обучить модель на большей части данных, сохраняя при этом независимые тестовые данные для оценки ее качества. Такое разделение помогает объективно оценить производительность модели, избегая переобучения. Обучающий набор данных используется для обучения различных базовых моделей, таких как метод опорных векторов, Random Forest, деревьев решений. Каждая модель обучается и изучает закономерности в данных, чтобы генерировать прогнозы в будущем. После используется процедура автоматического машинного обучения для автоматического выбора модели и модификации параметров. На этом этапе используются алгоритмы для определения оптимальной конфигурации путем перебора различных моделей и комбинаций параметров. Цель этой автоматизированной работы является эффективное улучшение производительности модели. Кроме того, используются методы ансамблирования, чтобы для учета прогнозов нескольких базовых моделей. Подходы, основанные на ансамбле, такие как классификаторы с голосованием, повышают точность прогнозирования за счет использования различных моделей [8].

*Метод опорных векторов.* В современных сетях является надежным методом, который умеет классифицировать и различать сетевые фрагменты на основе их уникальных свойств. Так как данный метод может работать со сложными данными и выявлять тонкие закономерности, такие как распределение полосы пропускания, ограничения по задержкам и качеству обслуживания. Метод опорных векторов находит подходящую гиперплоскость для разделения сетевых срезов, отображая входные данные в более высокоразмерное пространство, что позволяет точно идентифицировать и классифицировать их. Метод опорных векторов идеально подходит для распознавания сетевых срезов, помогает управлять и организовывать различными сетевыми ресурсами. Настройка модели была направлена на выбор ядра и оптимизацию гиперпараметров для достижения наилучших результатов. Поскольку SVM позволяет использовать различные функции ядра для преобразования данных в пространство более высокой размерности, были протестированы не-

сколько вариантов: линейное, полиномиальное и радиально-базисное (Radial Basis Function, RBF) ядра. Особым свойством метода опорных векторов является непрерывное уменьшение эмпирической ошибки классификации и увеличение зазора [9]. Наилучшие результаты были получены при использовании RBF-ядра, способного учитывать нелинейные зависимости в данных и эффективно разделять классы.

*Оптимизация гиперпараметров.* Для настройки модели использовались методы сеточного поиска (Grid Search) и случайного поиска (Randomized Search), позволяющие найти оптимальные значения гиперпараметров [10]. Параметр регуляризации  $C$  контролировал степень, до которой модель стремилась минимизировать ошибки. Высокое значение  $C$  усиливало точность разделения классов, но могло привести к переобучению. Параметр ядра  $\gamma$  (гамма) определял, насколько сильно отдельные обучающие примеры влияют на построение гиперплоскости. Высокое значение  $\gamma$  обеспечивало более детализированную границу разделения, тогда как низкое значение создавало более сглаженную границу. После завершения обучения производительность модели была испытана на тестовой выборке. Оценка включала расчет таких метрик, как точность и F1 Score, что позволило количественно оценить способность модели правильно классифицировать сетевые срезы. Модель SVM достигла точности 88,7% и значения F1 Score 0,87, что демонстрирует ее высокую точность при идентификации и классификации сетевых срезов. Далее в исследовании будет приведена таблица, в которой представлен результат точности.

Random Forest может управлять сложными и различными структурами сетевых данных, что делает его идеальным для обнаружения сетевых фрагментов. Random Forest хорошо работает, формируя ансамбль деревьев решений, которые оценивают и классифицируют сетевые срезы сети с помощью деревьев для классификации фрагментов сети на основе различных факторов. Путем объединения прогнозов нескольких деревьев решений, обученных на различных данных и характеристиках, Random Forest уменьшает избыточную подгонку и улучшает применимость модели к не встречающимся конфигурациям сетевых фрагментов. Выявление важнейших переменных и улавливание нелинейных взаимодействий между ними делает алгоритм идеальным для распознавания сложных моделей нарезки сети, что позволяет эффективно управлять и использовать различные сетевые ресурсы. Полученная точность достигает 91,5%.



*Дерево решений.* Деревья решений – это интуитивные и достаточно эффективные методы распознавания сетевых срезов, которые могут обрабатывать сложные взаимодействия атрибутов [11]. Деревья решений классифицируют сетевые срезы сети путем иерархического разбиения пространства признаков на основе значений атрибутов. Внутренние узлы дерева обозначают признаки и границы принятия решений в то время, как листовые узлы представляют конечные категории или классы. Деревья решений помогают объяснить распознавание сетевых срезов благодаря ясности логики классификации. Деревья решений могут работать с числовыми и категориальными данными и легко интерпретируются, что делает их полезными для идентификации и классификации фрагментов сети по их различным взаимосвязанным качествам. Полученная точность достигает 91,4% и F1 Score 0,90.

### Оценка моделей

Показатели точности используются для оценки способности модели предсказывать срез сети. Общая точность измеряет, насколько часто предсказания оказываются правильными. Однако метрика оценки качества классификационных алгоритмов уравнивает точность. Это помогает при несбалансированности классов. Такое сочетание позволяет провести полное исследование, гарантирующее, что модель может надежно обнаруживать ключевые события и включать все значимые фрагменты сети. Для определения метрики оценки качества классификационных алгоритмов применяется формула:

$$F1 = 2 * \frac{precision * recall}{precision + recall}. \quad (1)$$

Этот показатель помогает модели сбалансировать точность и полноту, минимизируя влияние ложных положительных и ложных отрицательных предсказаний. Чем выше F1-score, тем лучше модель управляет балансом между точностью и полнотой, что важно для точного распознавания и классификации сетевых срезов.

F1-score – это гармоническое среднее между точностью (Precision) и полнотой (Recall), которые измеряют способность модели правильно определять ключевые события и покрывать все значимые классы. F1-score особенно важен при дисбалансе классов, так как помогает уравновесить точность предсказаний и полноту, обеспечивая сбалансированную оценку.

$$precision = \frac{TP}{TP + FP}. \quad (2)$$

$$recall = \frac{TP}{TP + FN}. \quad (3)$$

Формула (1) позволяет рассчитать F1 Score. Формулы (2) и (3) используются для расчета формулы (1).

Precision (точность) – доля правильных предсказаний среди всех положительных предсказаний, отражающая способность модели избегать ложных положительных предсказаний.

Recall (полнота) – доля правильно классифицированных положительных примеров среди всех положительных примеров в наборе данных. Позволяет определить долю правильных прогнозов по сравнению с общим числом прогнозов. Оценка общей точности модели категоризации является важным показателем. Матрица несоответствий – это табличное представление, которое часто используется в области машинного обучения для оценки эффективности системы классификации [12]. В таблице 4 отображается количество истинно положительных, истинно отрицательных и ложных результатов, где TP – True Positive (Истинно позитивные), FP – False Positive (Ошибочно позитивные), FN – False Negative (Ошибочно негативные), TN – True Negative (Истинно негативные).

Таблица 4. Основные характеристики имитационной модели

|          | Positive | Negative |
|----------|----------|----------|
| Positive | TP       | FP       |
| Negative | FN       | TN       |

### Результаты оценки

Базовые модели, которые использовались, состоят из классификаторов, включая SVM, RF, дерево решений. Используя обучающий набор данных, обучили различные модели и оценили их производительность на основе точности и F1. Согласно информации, представленной в таблице 5, можно сделать вывод, что модель RF является наиболее эффективной для данного набора данных с точностью 91,5% и F1 – 0,902 балла.

Таблица 5. Основные характеристики имитационной модели

| Алгоритм       | Точность | F1    |
|----------------|----------|-------|
| SVM            | 0,887    | 0,870 |
| RF             | 0,915    | 0,902 |
| Дерево решений | 0,914    | 0,900 |

Далее, применяя базовые модели и технику настройки параметров к каждой модели оцениваем, повышается ли точность и F1 моделей. Согласно таблице 6, точность каждой модели заметно уве-

личилась, что повышает производительность и надежность классификации срезов сети.

Таблица 6. Основные характеристики имитационной модели

| Алгоритм       | Точность | F1    | Подбор лучшего параметра |
|----------------|----------|-------|--------------------------|
| SVM            | 0,887    | 0,870 | Kernel – liner           |
| RF             | 0,915    | 0,902 | N_estimation – 10        |
| Дерево решений | 0,914    | 0,900 | Max_depth – None         |

## Заключение

Сетевая нарезка имеет широкий спектр потенциальных применений в будущем, включая инновации в управлении сетью и усовершенствования. Адаптивные алгоритмы, которые могут динамически изменять оптимизацию сетевых срезов в ответ на фактические условия сети. Сетевые срезы позиционируются как жизненно важный инструмент для защиты будущего сетевых инфраструктур благодаря своей способности адаптироваться к изменяющимся типам сетевых дизайнов и грядущим технологиям, таким как 5G и пр. Использование алгоритмов обнаружения аномалий в сети подчеркивает значимость сетевых срезов в повышении отказоустойчивости и безопасности, гарантируя своевременное выявление и устранение возможных проблем. Кроме того, использование искусственного интеллекта и машинного обучения для предиктивной аналитики способно полностью изменить распределение сетевых ресурсов, что обеспечит получение лучшей и более эффективной сетевой нарезки. Изучение стратегий межслойной оптимизации может также обеспечить большую область знаний в управлении сетевыми ресурсами, позволяя адаптировать услуги к потребностям пользователей и пропускной способности сети. По сути, телекоммуникационную отрасль ожидает трансформация в результате роста возможностей сетевой нарезки, что обеспечит появление более надежных, безопасных и эффективных сетей.

## Литература

- Shen X. Device-to-device communication in 5G cellular networks // IEEE Network. 2015. Vol. 29, no. 2. P. 2–3.
- Flexible multi-node simulation of cellular mobile communications: the Vienna 5G System Level Simulator / M.K. Muller [et al.] // EURASIP Journal on Wireless Communications and Networking. 2018. URL: <https://jwcn-urasipjournals.springeropen.com/articles/10.1186/s13638-018-1238-7> (дата обращения: 09.11.2024).
- End-to-end simulation of 5G mmwave networks / M. Mezzavilla [et al.] // IEEE Communications Surveys & Tutorials. 2018. Vol. 20, no. 3. P. 2237–2263.
- Peuster M., Karl H., Rossem S. Medicine: Rapid prototyping of production-ready network services in multi-pop environments // 2016 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN). Palo Alto, 2016. DOI: 10.1109/NFV-SDN.2016.7919490
- Отбор признаков в задачах машинного обучения. Часть 1. URL: <https://habr.com/ru/articles/550978/> (дата обращения: 09.10.2024).
- F1 Score vs ROC AUC vs Accuracy vs PR AUC: Which evaluation metric should you choose? URL: <https://neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc> (дата обращения: 09.10.2024).
- Generalized hybrid beamforming for vehicular connectivity using THZ massive MIMO / S.A. Busari [et al.] // IEEE Transactions on Vehicular Technology. 2019. Vol. 68, no. 9. P. 8372–8383.
- Ансамблевые методы машинного обучения. URL: <https://habr.com/ru/articles/571296/> (дата обращения: 09.10.2024).
- Тарасов В.Н., Самаркин М.Е. Классификация больших данных телекоммуникационной компании с помощью технологии Data Mining // Инфокоммуникационные технологии. 2016. Т. 14, № 3. С. 258–263.
- Latif O.A., Amer M., Kwasinski A. Classification of network slicing requests using support vector machine // 2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA). Ras Al Khaimah, 2022. DOI: 10.1109/ICECTA57148.2022.9990459
- Hypergraph game theoretic solutions for load aware dynamic access of ultra-dense small cell networks / X. Zhu [et al.] // KSII Transactions on Internet and Information Systems. 2018. Vol. 13, no. 2. DOI: 10.3837/tiis.2019.02.002
- Переживаемая Матрицу Несогласий – Confusion matrix. URL: <https://habr.com/ru/articles/758510> (дата обращения: 09.10.2024).

Дата поступления 18.11.2024

Дата принятия 18.12.2024

**Кузнецов Александр Андреевич**, аспирант кафедры управления в технических системах (УТС) Поволжского государственного университета телекоммуникаций и информатики (ПГУТИ). 443090, Российская Федерация, г. Самара, Московское шоссе, 77. Тел. +7 996 728-08-21. E-mail: a1997@internet.ru

**Тарасов Вениамин Николаевич**, д.т.н., профессор, заведующий кафедрой УТС ПГУТИ. 443090, Российская Федерация, г. Самара, Московское шоссе, 77. Тел. +7 846 228-00-13. E-mail: v.tarasov@psuti.ru

UDC 004.89

DOI: 10.18469/ikt.2024.22.3.05

## RESEARCH OF MACHINE LEARNING ALGORITHMS FOR RECOGNIZING NETWORK SLICES

*Kuznetsov A.A., Tarasov V.N.*

*Povolzhskiy State University of Telecommunications and Informatics, Samara, Russian Federation*

*E-mail: a1997@internet.ru, v.tarasov@psuti.ru*

This article is devoted to the study of the application of machine learning algorithms for network slicing recognition within the framework of 5G telecommunication technologies. In the context of the dynamic development of network technologies, network segmentation is becoming a key tool for optimizing network services. The author analyzes various machine learning approaches and methods in order to effectively identify and classify network slicing, which is critical issue for proper management of the network resources and performance improvement of in general. The purpose of the study is to explore models in order to find one that can accurately identify the most appropriate network fragments for various services, which, in result, contributes to improved network performance and more efficient use of resources. Results and conclusions of the study can be useful for development of more efficient and automated network management methods based on network slices analysis using machine learning.

**Keywords:** *network slices, network slicing, 5G technologies, machine learning, network segmentation*

**Kuznetsov Alexander Andreevich**, Povolzhskiy State University of Telecommunications and Informatics, 77, Moskovskoe shosse, Samara, 443090, Russian Federation; Student of Control in Technical Systems Department. Tel. +7 996 728-08-21. E-mail: a1997@internet.ru

**Tarasov Veniamin Nikolaevich**, Povolzhskiy State University of Telecommunications and Informatics, 77, Moskovskoe shosse, Samara, 443090, Russian Federation; Head of Control in Technical Systems Department, Doctor of Technical Sciences, Professor. Tel. + 7 846 228-00-13. E-mail: v.tarasov@psuti.ru

### References

1. Shen X. Device-to-device communication in 5G cellular networks. *IEEE Network*, 2015, vol. 29, no. 2, pp. 2–3.
2. Muller M.K. et al. Flexible multi-node simulation of cellular mobile communications: the Vienna 5G System Level Simulator. *EURASIP Journal on Wireless Communications and Networking*, 2018. URL: <https://jwcn-urasipjournals.springeropen.com/articles/10.1186/s13638-018-1238-7> (accessed: 09.11.2024).
3. Mezzavilla M. et al. End-to-end simulation of 5G mmwave networks. *IEEE Communications Surveys & Tutorials*, 2018, vol. 20, no. 3, pp. 2237–2263.
4. Peuster M., Karl H., Rossem S. Medicine: Rapid prototyping of production-ready network services in multi-pop environments. *2016 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*. Palo Alto, 2016. DOI: 10.1109/NFV-SDN.2016.7919490
5. Feature selection in machine learning problems. Part 1. URL: <https://habr.com/ru/articles/550978/> (accessed: 09.10.2024).

6. F1 Score vs ROC AUC vs Accuracy vs PR AUC: Which evaluation metric should you choose? URL: <https://neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc> (accessed: 09.10.2024).
7. Busari S.A. et al. Generalized hybrid beamforming for vehicular connectivity using THZ massive MIMO. *IEEE Transactions on Vehicular Technology*, 2019, vol. 68, no. 9, pp. 8372–8383.
8. Ensemble methods of machine learning. URL: <https://habr.com/ru/articles/571296/> (accessed: 09.10.2024).
9. Tarasov V.N., Samarkin M.E. Telecommunication company big data classification by data mining technique. *Infokommunikacionnye tekhnologii*, 2016, vol. 14, no 3, pp. 258–263. (In Russ.)
10. Latif O.A., Amer M., Kwasinski A. Classification of network slicing requests using support vector machine. *2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*. Ras Al Khaimah, 2022. DOI: 10.1109/ICECTA57148.2022.9990459
11. Zhu X. et al. Hypergraph game theoretic solutions for load aware dynamic access of ultra-dense small cell networks. *KSI Transactions on Internet and Information Systems*, 2018, vol. 13, no. 2. DOI: 10.3837/tiis.2019.02.002
12. Chewing the matrix of inconsistencies – confusion matrix. URL: <https://habr.com/ru/articles/758510/> (accessed: 09.10.2024). (In Russ.)

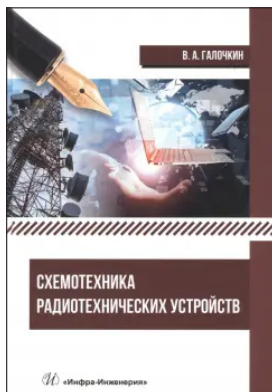
Received 18.11.2024

Accepted 18.12.2024

## Реклама

Галочкин, В.А.

**Схемотехника радиотехнических устройств** / В.А. Галочкин. – Москва; Вологда: Инфа-Инженерия, 2023. – 332 с.



ISBN 978-5-9729-1470-8

УДК 621.38  
ББК 32.844.1

Рассмотрены основные вопросы организации и построения систем и линий радиосвязи, радиовещания и телевидения, сотовые сети радиосвязи стандарта GSM, радиосвязь через Интернет, системы GPS и ГЛОНАСС, беспроводная локальная сеть Wi-Fi, стандарты беспроводных сетей Bluetooth, принципы радиолокации.

Исследуются принципы построения и схемотехника базовых узлов радиотехнических устройств: входных цепей, усилителей радиочастоты, амплитудных, частотных и фазовых детекторов, усилителей низкой частоты, систем авторегулирования усиления, систем автоматической подстройки частоты и др. Приведены примеры расчетов устройств, а также примеры их схемотехнического компьютерного моделирования.

Для студентов телекоммуникационных и радиотехнических специальностей дневной и заочной формы обучения, а также для инженерно-технических работников, изучающих электронику и схемотехнику.