

DOI: <https://doi.org/10.30898/1684-1719.2026.8.7>

УДК: 621.371.3 + 519.246

ИДЕНТИФИКАЦИЯ ПАРЦИАЛЬНЫХ КОМПОНЕНТ ИМПУЛЬСНОЙ ХАРАКТЕРИСТИКИ РАДИОКАНАЛА ПО ДАННЫМ ТРАССИРОВКИ ЛУЧЕЙ РЕАЛЬНОЙ ГОРОДСКОЙ ЗАСТРОЙКИ

А.В. Епонешников, А.И. Сулимов

Казанский (Приволжский) федеральный университет, Институт физики,
420008, Россия, Казань, ул. Кремлевская, д. 16а

Статья поступила в редакцию 1 июня 2026 г.

Аннотация. Рассмотрена проблема автоматической идентификации парциальных компонент импульсной характеристики радиоканала (ИХ) по данным трассировки лучей реальной городской застройки. Предложенный метод основан на комбинации корреляционной обработки зондирующих сигналов с дальномерными кодами Голда, вейвлет-фильтрации и классификации обнаруженных откликов алгоритмом градиентного бустинга XGBoost. Путем имитационного моделирования показано, что вероятность ложной идентификации удалось снизить до 0,8-3,4 %. Установлено, что основными факторами необнаружения лучей являются ошибки классификатора XGBoost (42-45 %), маскирование слабых компонент боковыми лепестками АКФ (26-28 %) и недостаточное временное разрешение близких лучей (22-23 %). Для задач формирования согласованных случайных последовательностей рекомендуется использовать 2-4 доминирующих луча: такой выбор обеспечивает охват 66-69 % мощности канала при вероятности ложного обнаружения 1,2-2,0 %, что соответствует требованиям задач физического уровня безопасности.

Ключевые слова: импульсная характеристика канала, многолучевое распространение, идентификация сигнала, трассировка лучей, Sionna RT, TDL,

3GPP TR 38.901, XGBoost, градиентный бустинг, коды Голда, вейвлет-шумоподавление.

Финансирование: Работа выполнена за счет средств Программы стратегического академического лидерства Казанского (Приволжского) федерального университета.

Автор для переписки: Епонешников Александр Вячеславович,
eponeshnikov.sanek@mail.ru

Введение

Идентификация импульсной характеристики радиоканала (ИХ) занимает важное место в технологиях современных систем беспроводной связи, в том числе в системах обеспечения информационной безопасности на физическом уровне (Physical Layer Security, PLS) [1, 2, 22]. Одним из перспективных направлений PLS является формирование согласованных случайных последовательностей на основе ИХ канала [3, 4]. Принцип основан на свойстве взаимности канала: абоненты А и В, расположенные на концах радиолинии, наблюдают одну и ту же многолучевую структуру канала, которая является источником общей случайности, недоступной третьей стороне. Случайная последовательность вырабатывается из задержек парциальных компонент канала.

Ключевое требование к таким системам – согласованная идентификация компонент ИХ на обоих концах радиолинии, что обеспечивает одинаковые оценки, информативных задержек лучей. Каждая ошибка идентификации приводит к рассогласованию формируемых последовательностей: ложно обнаруженная компонента (ошибка I рода) на одной стороне не будет иметь соответствия на другой, а пропущенная истинная компонента (ошибка II рода) уменьшает длину генерируемой абонентами последовательности. Поэтому минимизация вероятности ложного обнаружения FP является первоочередной задачей, определяющей практическую применимость метода.

Ранее авторами разработана методика автоматической идентификации компонент ИХ, основанная на зондировании канала с помощью дальномерных

кодов Голда, вейвлет-фильтрации принятого сигнала и классификации обнаруженных откликов алгоритмом градиентного бустинга XGBoost [6]. Методика проверена на стохастических канальных моделях TDL (Tapped Delay Line) стандарта 3GPP TR 38.901 [7, 11] с анализом более 2,7 миллионов парциальных компонент. В [5] исследована оценка энтропии ИХ гистограммным методом, что определило необходимость точной идентификации парциальных компонент как предварительного этапа PLS. Однако стохастические модели генерируют задержки и амплитуды из абстрактных статистических распределений и не учитывают геометрию реальной среды – архитектуру зданий, ширину улиц, высоту размещения передатчика.

Для устранения этого недостатка в данной статье разработанная методика идентификации компонент ИХ будет опробована на результатах трассировки лучей трех реальных городских сцен различного типа: плотная застройка (Мюнхен), низкоэтажная застройка (Флоренция) и квартальная сетка с высотной застройкой (Сан-Франциско). Трассировка лучей выполнена в библиотеке Sionna RT [10], реализующей дифференцируемый трассировщик с ускорением на графических процессорах.

Целью данной работы является оценка эффективности идентификации парциальных компонент ИХ с учетом реалистичной геометрии городской застройки и проверка возможности обучения классификатора на синтетических данных без потери качества его работы применительно к задачам PLS.

Для достижения цели решены следующие задачи: I. Проверка успешности работы классификатора для решения задач PLS; II. Рекомендации по использованию ИХ для PLS; III. Факторы ошибочной классификации; IV Влияние геометрии среды.

1. Методы идентификации компонент импульсной характеристики

ИХ многолучевого канала описывается суммой парциальных компонент, каждая из которых характеризуется задержкой τ_i , амплитудой a_i и фазой φ_i .

Импульсная характеристика канала описывается выражением (1):

$$h(\tau) = \sum_{i=1}^n a_i \cdot e^{j\varphi_i} \cdot \delta(\tau - \tau_i), \quad (1)$$

где n – число парциальных компонент. Задача идентификации состоит в определении числа компонент n , их задержек $\{\tau_1, \tau_2, \dots, \tau_n\}$ и амплитуд $\{a_1, a_2, \dots, a_n\}$ по принятому зашумленному сигналу. Для формирования случайных последовательностей (СП) в системах PLS к методу идентификации предъявляются особые требования: 1) работоспособность при низком отношении сигнал/шум (SNR); 2) высокое временное разрешение для разделения близких компонент; 3) минимизация ложных обнаружений (каждое из которых приводит к рассогласованию формируемых последовательностей в A , B); 4) отсутствие строгих априорных предположений о структуре канала.

Методы сверхразрешения обеспечивают точную оценку задержек за пределами инструментального разрешения. Итерационный алгоритм наименьших квадратов LIMS [12] последовательно вычитает из корреляционной функции вклад уже обнаруженных компонент и повторно оценивает оставшиеся методом наименьших квадратов, что позволяет разделять лучи без полного пересчета всех параметров на каждом шаге. Метод сверхразрешения на основе взвешенной реконструкции [13] дополнительно подавляет боковые лепестки автокорреляционной функции (АКФ), разрешая компоненты с задержками менее длительности чипа. Однако оба метода требуют априорного знания числа лучей n , а вычислительная сложность растет как $O(n^2)$, что делает их непригодными для массовой обработки тысяч реализаций канала. Метод идентификации значимых компонент на основе сингулярного спектрального анализа [17] разделяет детерминированные составляющие и шум без знания n , однако требует множественных наблюдений одного и того же канала (серии OFDM-символов), что ограничивает его применимость системами с повторным доступом к каналу.

При корреляционной обработке широкополосных зондирующих сигналов помимо полезных компонент возникает шумовая дорожка – смесь боковых

лепестков АКФ и аддитивного шума, на фоне которой слабые компоненты ИХ неразличимы. Для подавления шумовой дорожки без искажения формы полезных компонент применяется вейвлет-шумоподавление. Метод VisuShrink [24] разлагает АКФ отклика канала на вейвлет-коэффициенты различных масштабов и подавляет мелкомасштабные коэффициенты, не превышающие порог $\lambda = \sigma\sqrt{2\ln N}$, где σ – оценка стандартного отклонения шума. Реконструированный после фильтрации сигнал сохраняет форму откликов истинных компонент ИХ, но вместе с шумовой дорожкой подавляет и слабые компоненты. Тем не менее, вейвлет-шумоподавление не решает задачу классификации: среди оставшихся откликов присутствуют как истинные компоненты, так и остаточные артефакты обработки, что требует последующего разделения.

Методы машинного обучения (ML) активно применяются для оценки различных параметров радиоканала: позиционирования по ИХ [21], классификации условий распространения LOS/NLOS [15], оценки канала в системах 5G/6G [14]. Мы используем ML для классификации компонент АКФ на истинные лучи и артефакты. Наибольшую эффективность показал градиентный бустинг по методу XGBoost [6], использующему ансамбль решающих деревьев, последовательно минимизирующий ошибку предыдущих деревьев. В задаче распознавания LOS/NLOS-состояний канала, XGBoost достиг точности 82,8 % на наборе из 42000 образцов, превзойдя такие методы как KNN, MLP и SVM [15].

Для интерпретации решений классификатора используется метод SHAP (SHapley Additive exPlanations) [16], основанный на теории кооперативных игр Шепли. SHAP присваивает каждому признаку числовое значение его вклада в конкретное предсказание модели, что позволяет контролировать физическую обоснованность решений: например, убедиться, что классификатор опирается на выраженность отклика, а не на артефактные корреляции. Сочетание XGBoost и SHAP обеспечивает прозрачную классификацию, выгодно отличающую данный подход от нейросетевых методов, работающих по принципу «черного ящика» [14, 21]. Вместе с тем ML-методы зависят от полноты обучающих

данных: модель, обученная на одном типе каналов, может неадекватно работать на другом, что вызывает проблему подготовки корректного массива обучающих данных.

Описанные ограничения каждой группы методов не позволяют обособленно использовать их для задачи формирования случайных последовательностей (PLS). В настоящей работе используется комбинированный подход, последовательно устраняющий эти ограничения. Зондирование канала с помощью дальномерных кодов обеспечивает выигрыш обработки ≈ 33 дБ, что снимает как необходимость в априорной информации о числе компонент (ограничение методов сверхразрешения), так и требование к высокому SNR (ограничение спектрального анализа). Вейвлет-шумоподавление подавляет шум и боковые лепестки АКФ без искажения формы откликов, но не различает истинные компоненты и артефакты. Последнее решается методом XGBoost [6] с возможностью контроля правильности классификации через SHAP-анализ. Открытым остается вопрос о необходимости обучения классификатора на реалистичных данных, учитывающих геометрию реальной среды распространения, что подразумевает трудоемкую процедуру подготовки ее детальной модели.

2. Модель системы

2.1. Модель канала

Стандарт 3GPP TR 38.901 [11] определяет набор стохастических моделей канала с линией задержки с отводами (Tapped Delay Line, TDL) для диапазона частот 0,5-100 ГГц. Модели TDL получаются из кластерных моделей CDL (Clustered Delay Line) путем пространственной фильтрации с учетом диаграммы направленности антенны. Каждая модель задает фиксированный набор отводов с нормированными задержками и мощностями; при моделировании задержки масштабируются в соответствии с заданным среднеквадратическим разбросом задержек (Delay Spread, DS). Модели TDL-A, TDL-B и TDL-C описывают условия без прямой видимости (NLOS), модели TDL-D

и TDL-E – условия прямой видимости (LOS) с К-фактором Райса. Достоинство TDL-моделей – вычислительная эффективность и простота параметризации; ограничение – невозможность предсказать канал для конкретной пространственной конфигурации. Параметры моделей приведены в табл. 2.

Детерминистическое моделирование выполнено с помощью библиотеки Sionna RT [10], реализующей методы геометрической оптики [8] для городских сцен 5G [9]. Для каждой сцены задана трехмерная модель городской среды, положения передатчиков (TX) на крышах зданий и приемников (RX) на уровне улицы. Трассировщик рассчитывает пути лучей с учетом зеркальных отражений и преломления до глубины 5 взаимодействий. Статистика замираний в таких каналах описывается логнормальными и рэлеевскими моделями [18, 19], параметры которых зависят от плотности застройки [20]. Используются три городские сцены различного типа, параметры которых приведены в табл. 1.

Таблица 1. Параметры RT-сценариев (Sionna RT).

Параметр	Мюнхен	Флоренция	Сан-Франциско
Тип сцены	UMi, плотная	UMi, низкоэтажная	UMa, кварт. сетка
Глубина трассировки	5	5	5
Ср. расст. радиолинии, м	347	253	529
Медиана расст. радиолинии, м	303	218	515
Число радиолиний	2400	2400	2400

Таблица 2. Параметры стохастических TDL-моделей (3GPP TR 38.901).

Параметр	TDL-A	TDL-B	TDL-C	TDL-D	TDL-E
Условия	NLOS	NLOS	NLOS	LOS	LOS
Число отводов	23	23	24	13	14
К-фактор, дБ	–	–	–	13,3	22
DS (масштаб.), нс	100; 300; 600	100; 300; 600	100; 300; 600	100; 300; 600	100; 300; 600
Фединг	Рэлей	Рэлей	Рэлей	Рэлей/Райс	Рэлей/Райс

2.2. Модель зондирующего сигнала

Зондирование канала выполняется радиосигналом, несущую которого модулирует псевдослучайная последовательность – код Голда длиной $N = 2047$ чипов [23]. Код Голда формируется поразрядным сложением

по модулю 2 двух М-последовательностей, порожденных примитивными полиномами степени 11. Бинарная последовательность кода модулирует несущую методом двоичной фазовой манипуляции (Binary Phase Shift Keying, BPSK).

Модулированный сигнал проходит через формирующий фильтр типа «квадратный корень из приподнятого косинуса» (Root Raised Cosine, RRC) с коэффициентом сглаживания $\beta = 0,22$ при символьной скорости $R_s = 200$ МГц. Применение согласованного фильтра на приемной стороне обеспечивает выигрыш обработки (processing gain) $PG = 10 \lg N \approx 33$ дБ.

Таблица 3. Параметры зондирующего сигнала.

Параметр	Значение
Тип зондирующего кода	Код Голда
Длина кода	2047 чипов
Модуляция	BPSK
Формирующий фильтр	RRC, $\beta = 0,22$
Символьная скорость	200 МГц
Диапазон SNR	-10, -5, 0, 5, 10, 20 дБ

2.3. Методы обработки сигнала

На приемной стороне выполняется согласованная фильтрация: вычисляется корреляция принятого сигнала $r(t)$ с копией зондирующего сигнала $s(t)$ согласно выражению (2):

$$R(\tau) = \int r(t) \cdot s(t - \tau) dt, \quad (2)$$

где $s(t)$ – комплексно-сопряженный зондирующий сигнал. Корреляционная функция $R(\tau)$ содержит отклики в точках задержек парциальных компонент. Для подавления шумовых откликов применяется вейвлет-шумоподавление по методу VisuShrink [24] с универсальным порогом $\lambda = \sigma \sqrt{2 \ln N}$, где σ – оценка стандартного отклонения шума.

Обнаруженные отклики классифицируются моделью градиентного бустинга XGBoost [6] по двум классам: истинная компонента многолучевого канала или артефакт обработки (шумовой отклик, боковой лепесток).

Для каждого отклика формируется вектор признаков трех уровней: (а) локальные признаки компоненты (10) – высота, выраженность, ширина на полувысоте, пороги и положения оснований; (б) контекстные признаки – аналогичные характеристики трех ближайших компонент слева и справа; (в) глобальные статистики – среднее, медиана, минимум и максимум локальных характеристик по всем обнаруженным компонентам данной реализации.

2.4. Метрики качества работы системы

Для задачи формирования согласованных последовательностей решающее значение имеют две метрики: 1) вероятность ложного обнаружения (FP, %) – отношение числа ложно обнаруженных компонент n_F к общему числу обнаруженных n_D ; 2) вероятность правильного распознавания компонент многолучевого кластера (TP, %) – отношение числа правильно идентифицированных компонент n_T к общему числу истинных n . Каждая ложно обнаруженная компонента на одной стороне радиолинии не имеет соответствия на другой и приводит к рассогласованию формируемой последовательности, поэтому минимизация FP является первоочередной задачей. В свою очередь, вероятность правильного распознавания определяет длину генерируемой последовательности.

Каждый луч из эталонной модели канала получает одну из категорий: (а) успешно идентифицирован (TP); (б) потеря по маскированию; (в) потеря по временному разрешению; (г) ошибка классификатора (FN); (д) прочие потери.

Для оценки практической значимости распознанных компонент введен параметр N – число используемых сильнейших по амплитуде компонент: из всех откликов, классифицированных как истинные, выбираются N сильнейших, остальные отбрасываются. Это моделирует практический сценарий, в котором для формирования последовательности используются лишь наиболее надежные компоненты, а число ложных обнаружений ограничивается сверху. Ключевые параметры моделирования сведены в табл. 3.

Все метрики вычислены для компонент многолучевого кластера за исключением луча прямой видимости (LoS): прямой луч обнаруживается

тривиально как сильнейшая компонента и не является источником случайности для формирования согласованных последовательностей.

3. Характеристика набора данных

В рамках моделирования для каждой городской сцены были определены 12 наиболее высотных зданий, на крыше каждого из них размещен передатчик. Далее на уровне улицы было случайно размещены 200 приемников, что дает 2400 случайно ориентированных радиолиний на сцену (7200 радиолиний суммарно для трех городских сцен). На рис. 1 представлено эмпирическое распределение длины модельных радиолиний для объединенного набора данных всех городских сцен. Медианная длина составила 312 м, средняя – 371 м. Распределение оказалось близко к логнормальному, что согласуется с общепринятой моделью [25]. Каждый из рассмотренных городов обладал уникальной архитектурой и геометрией застройки, что предопределило вариацию статистических характеристик городских сцен (см. табл. 1).

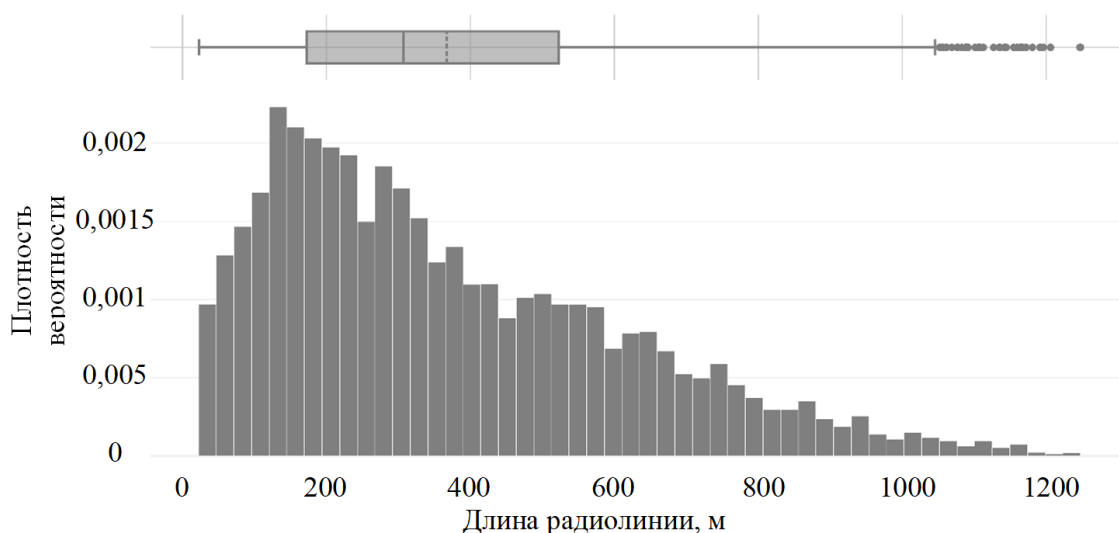


Рис. 1. Распределение расстояний радиолиний (все RT-сцены).

На рис. 2 показано эмпирическое распределение высоты передающей антенны модельных радиолиний. Во Флоренции преобладали малоэтажные здания высотой 20-30 м, в Мюнхене – застройка промежуточного типа с высотами 30-60 м, в Сан-Франциско – преобладали высотные здания (120-170 м).

Разнообразие высот подвеса антенн и длин радиолиний обеспечило требуемую репрезентативность условий распространения сигнала.

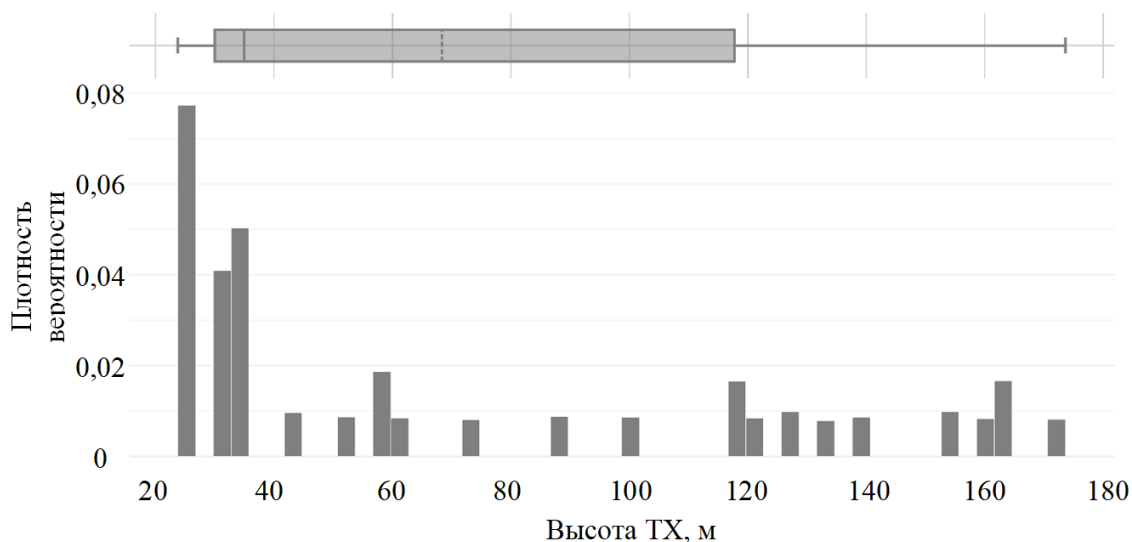


Рис. 2. Распределение высот передатчиков (все RT-сцены).

Для каждой радиолинии методом RT были рассчитаны ИХ при 6 значениях SNR. Далее при сопоставимых параметрах застройки и аналогичных SNR генерировались стохастические ИХ с использованием статистических моделей TDL-A...TDL-E. При этом тип модели канала определял среднее количество компонент многолучевого кластера (без учета LoS): в TDL-моделях оно фиксировано стандартом, в RT-моделях зависит от геометрии среды (см. табл. 4). Характерно, что среднее число лучей в RT-модели оказалось в 5-6 раз меньше, чем предполагалось синтетическими моделями TDL. С другой стороны, большинство отводов TDL имеют низкую мощность и не оказывают значимого влияния на структуру принимаемого сигнала.

Таким образом, сформированный набор данных (7200 радиолиний, три городских сцены с различной геометрией застройки, шесть уровней SNR) охватывает широкий диапазон условий распространения – от плотной малоэтажной застройки до высотных кварталов – и обеспечивает репрезентативность результатов относительно реальных городских сценариев.

Таблица 4. Среднее число компонент многолучевого кластера.

Сценарий / Модель	Тип	Ср. в кластере
Мюнхен	RT	5,57
Флоренция	RT	5,53
Сан-Франциско	RT	3,27
TDL-A (NLOS)	TDL	23 (фикс.)
TDL-B (NLOS)	TDL	23 (фикс.)
TDL-C (NLOS)	TDL	24 (фикс.)
TDL-D (LOS)	TDL	12 (фикс.)
TDL-E (LOS)	TDL	13 (фикс.)

4. Результаты

Для оценки качества идентификации компонент многолучевого кластера проведены имитационные эксперименты для четырех комбинаций набора «обучающих-тестовых» данных: RT→RT, TDL→TDL, RT→TDL, TDL→RT. В первых двух случаях обучение и тестирование классификатора лучей проводилось на однородных данных, в последних двух случаях – на смешанных данных. Последний случай представлял особый интерес, поскольку возможность обучения классификатора на синтетических данных модели TDL значительно упрощает задачу его разработки, избавляя от рутинной необходимости длительной подготовки детализированных 3D-карт городской среды.

4.1. Анализ адекватности обучения классификатора

Диаграммы на рис. 3 показывают процент успешно распознанных лучей для различных комбинаций обучающих/тестовых данных при различных уровнях SNR. На рис. 3 и 4 столбцы 1-4 соответствуют комбинациям «обучение → тестирование»: 1 – RT→RT; 2 – TDL→RT; 3 – RT→TDL; 4 – TDL→TDL. Наибольший интерес представлял случай «TDL→RT», в котором обучение классификатора происходило по синтетическим данным абстрактной модели TDL, а тестирование – по расчетным ИХ реальной городской среды. Примечательно, что комбинация «TDL→RT» обеспечила более высокую вероятность распознавания (44-52 %), чем «RT→RT» (39-45 %). Обучение на многочисленных отводах TDL-модели сделало классификатор более чувствительным к слабым

компонентам. Однако эта чувствительность обернулась ростом ложных обнаружений (рис. 4), что нежелательно для задач генерации СП на основе ИХ. В частности, вероятность FP для данных «TDL→RT» достигала 10,5 % и не опускалась ниже 5,5 %.

Для наиболее корректной, но трудоемкой в осуществлении, комбинации данных «RT→RT» вероятность ложного обнаружения оказалась в 5-7 раз ниже и варьировалась в пределах 0,8-2,6 %. Комбинация данных «RT→TDL» подтвердила гипотезу о том, что процесс обучения на данных RT действует консервативным образом: классификатор становится более осторожным в принятии решений и игнорирует слабые лучи. Хотя при этом достигалась низкая вероятность ложного обнаружения $FP = 0,9-2,2 \%$, успешно распознавалось всего 22-28 % лучей. Сводные результаты приведены в табл. 5.

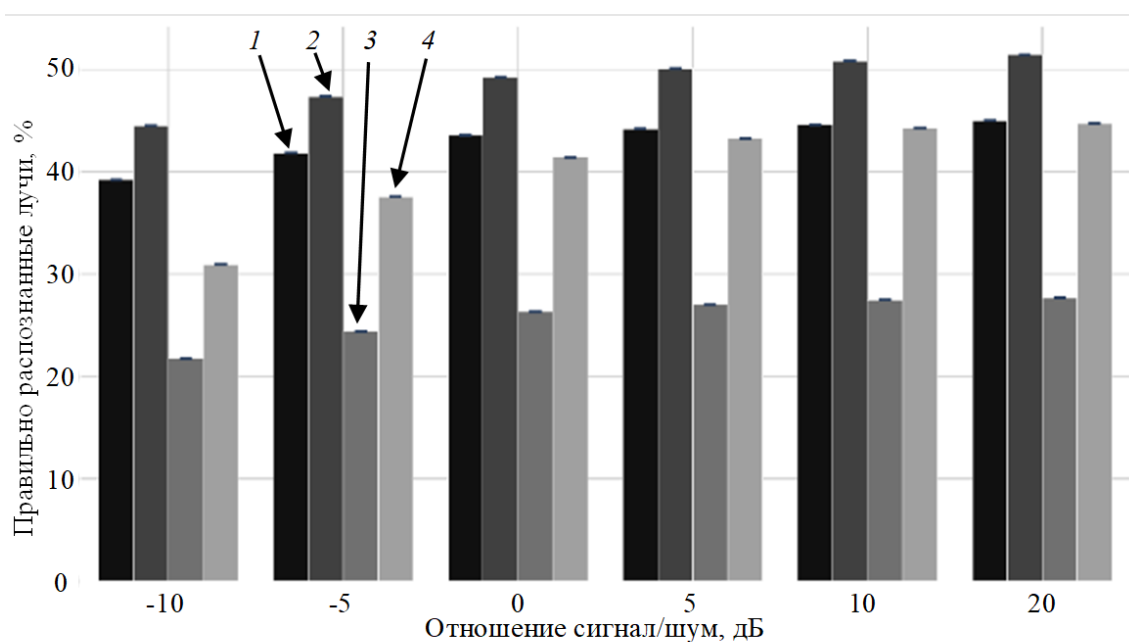


Рис. 3. Доля правильно распознанных компонент многолучевого кластера от SNR (все комбинации обучения/тестирования): 1 – RT→RT; 2 – TDL→RT; 3 – RT→TDL; 4 – TDL→TDL.

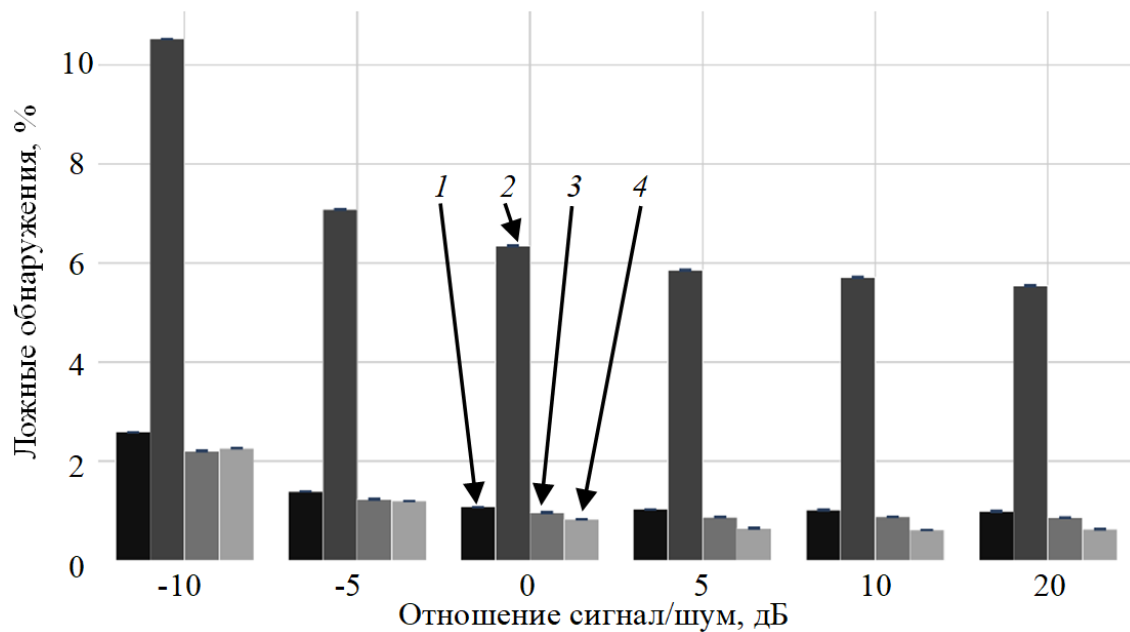


Рис. 4. Доля ложных обнаружений (FP) от SNR (все комбинации обучения/тестирования): 1 – RT→RT; 2 – TDL→RT; 3 – RT→TDL; 4 – TDL→TDL.

Таблица 5. Индикаторы производительности классификатора (SNR = 10 дБ).

Обучение → Тест	ТР, %	ФР, %	Расп., %
RT → RT	47,4	1,02	44,6
TDL → RT	54,5	5,71	50,8
RT → TDL	31,4	0,87	27,5
TDL → TDL	47,2	0,61	44,3

Возникает вопрос о компромиссе между чувствительностью и надежностью классификатора. Чтобы ответить на него, была рассмотрена суммарная доля мощности успешно обнаруженных лучей случайно рассеянного кластера. Введем необходимые обозначения. Пусть N_{Σ} – суммарное количество лучей в канале, случайно рассеянных на элементах среды. При этом пусть классификатор обнаруживает N_D лучей, причем, вообще говоря, $N_D \neq N_{\Sigma}$, так как не исключены ошибки I и II родов. Проранжируем обнаруженные лучи в порядке убывания мощности: $\{P_1, P_2, \dots, P_N, P_{N+1}, \dots, P_{N_D}\}$, где P_1 – максимальная мощность в рассеянном кластере, P_{N_D} – соответственно минимальная мощность. Выделим первые N наиболее мощных (главных) обнаруженных лучей. Пусть среди этих N главных парциальных компонент классификатор правильно распознал $N_I \leq N$, $N_I \leq N_{\Sigma}$ истинных лучей в канале.

На рис. 5 показано, какую долю мощности $E_h\{\sum_{j \in N_I} P_j / \sum_{i=1}^{N_\Sigma} P_i\}$ многолучевой

компоненты сигнала корректно обнаруживает классификатор среди N главных выявленных им компонент при усреднении по всему массиву 7200 импульсных характеристик, рассчитанных методом RT. Вместе с тем, на рис. 6 показана доля ложных обнаружений среди N главных обнаруженных классификатором компонент. На рисунках показано четыре кривые, соответствующие двум уровням $\text{SNR} = \{-10 \text{ дБ}; 20 \text{ дБ}\}$ и двум типам данных, использованных при обучении классификатора (RT/TDL). На рис. 5 и 6 кривые 1-4 соответствуют: 1 – RT, $\text{SNR} = -10 \text{ дБ}$; 2 – RT, $\text{SNR} = 20 \text{ дБ}$; 3 – TDL, $\text{SNR} = -10 \text{ дБ}$; 4 – TDL, $\text{SNR} = 20 \text{ дБ}$.

Из рис. 5 видно, что главная обнаруженная классификатором компонента ($N = 1$) в среднем объясняла свыше половины рассеянной мощности в канале. При этом из рис. 6 видно, что эта доминирующая в канале компонента успешно обнаруживалась классификатором не менее чем в 96 % случаях, а вероятность ошибочного обнаружения не превосходила 4 %. Добавление второй главной обнаруженной компоненты ($N = 2$) позволило успешно объяснить 66-69 % рассеянной мощности канала, однако процент ложных обнаружений возрос до ~6 %. Вторая компонента уже не столь уверенно выделялась на фоне шума, поэтому наблюдались заметные отличия в проценте ложных обнаружений для разных SNR. Начиная с $N = 3$, каждая последующая обнаруженная классификатором компонента давала лишь небольшой прирост в описании рассеянного кластера канала, но значительно ухудшала достоверность обнаружения.

Из рис. 5-6 следует, что консервативность классификатора при обучении на реалистичных данных RT более предпочтительна, чем высокая чувствительность при обучении на синтетических данных TDL. Как оказалось, нет большого смысла в обнаружении всех парциальных компонент канала, так как большинство из них имеет крайне низкую мощность и слабо применимы на практике для генерации СП. При этом высокая чувствительность системы порождает проблему ложных обнаружений, что особенно нежелательно

для целей генерации СП. В дальнейшем будет использоваться обучение классификатора только на данных трассировки лучей.

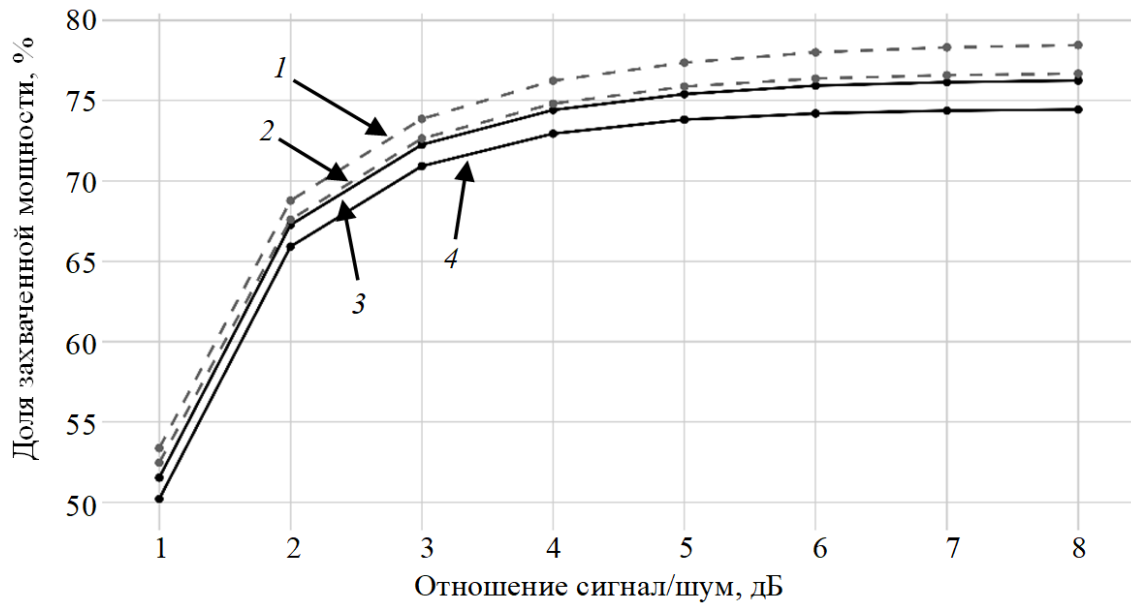


Рис. 5. Доля обнаруженной мощности – число главных обнаруженных компонент: 1 – TDL, SNR = 20 дБ; 2 – RT, SNR = 20 дБ; 3 – TDL, SNR = -10 дБ; 4 – RT, SNR = -10 дБ.

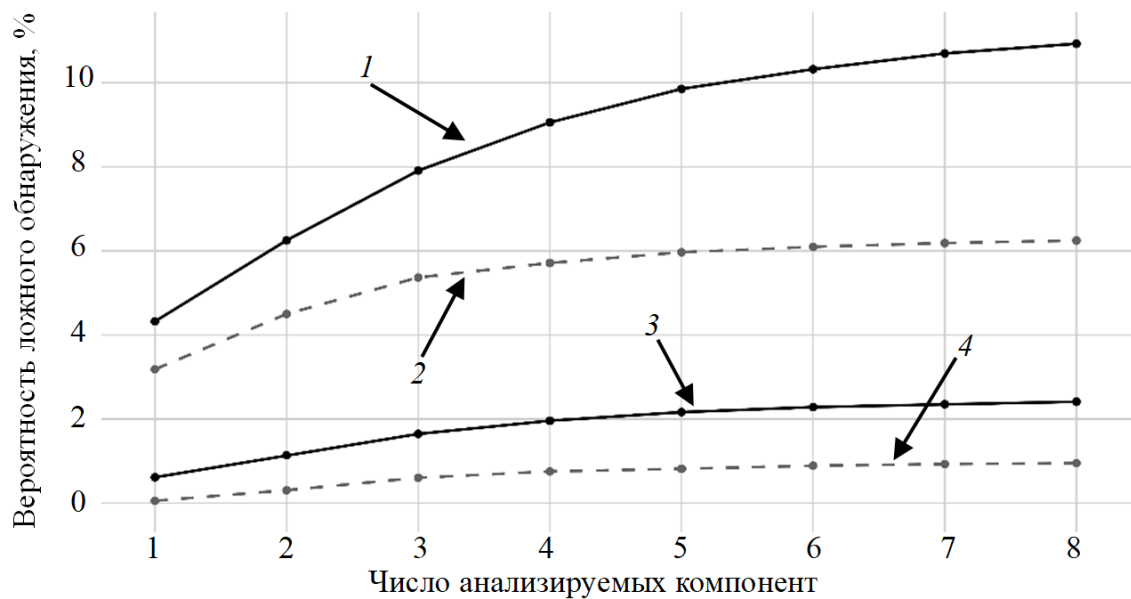


Рис. 6. Вероятность ложного обнаружения – число главных обнаруженных компонент: 1 – TDL, SNR = -10 дБ; 2 – TDL, SNR = 20 дБ; 3 – RT, SNR = -10 дБ; 4 – RT, SNR = 20 дБ.

4.2. Основные факторы ошибки второго рода

На рис. 7 представлено процентное соотношение различных причин, помешавших успешной идентификации лучей в канале. Сегменты столбцов 1-4 соответствуют: 1 – ошибка классификатора (FN); 2 – потеря по временному разрешению; 3 – потеря по маскированию; 4 – прочие потери. Основной причиной необнаружения лучей (FN, 42-45 % от суммарных потерь) оказались ошибки классификации. Зачастую отклик канала различался системой, однако модель машинного обучения XGBoost не распознавала в нем парциальный луч по набору наблюдаемых признаков.

К ошибкам второго рода примерно в 26-28 % случаев также приводит маскирование слабых компонент канала боковыми лепестками корреляционной функции мощных лучей. Ограничение исходит от длины дальномерного кода Голда в зондирующем сигнале (2047 чипов), для которого соотношение (сигнал/помеха) определяется выражением (3):

$$DR = 10 \cdot \lg(N / t(n)) = 10 \cdot \lg(2047 / 65) \approx 15 \text{ дБ}, \quad (3)$$

где $t(n) = 65$ – пиковый уровень боковых лепестков автокорреляционной функции (АКФ) кода [23]. Анализ данных показал, что, если мощность сильного луча на 17 дБ превосходила мощность слабого луча, расположенного по соседству в ИХ, то последний маскировался боковыми лепестками АКФ и не обнаруживался системой.

Наконец, в 22-23 % случаев сказывались потери по временному разрешению. Два луча с относительными задержками менее $1,5T$ (T – длительность чипа дальномерного кода) не разрешались системой, вследствие чего один из лучей не обнаруживался.

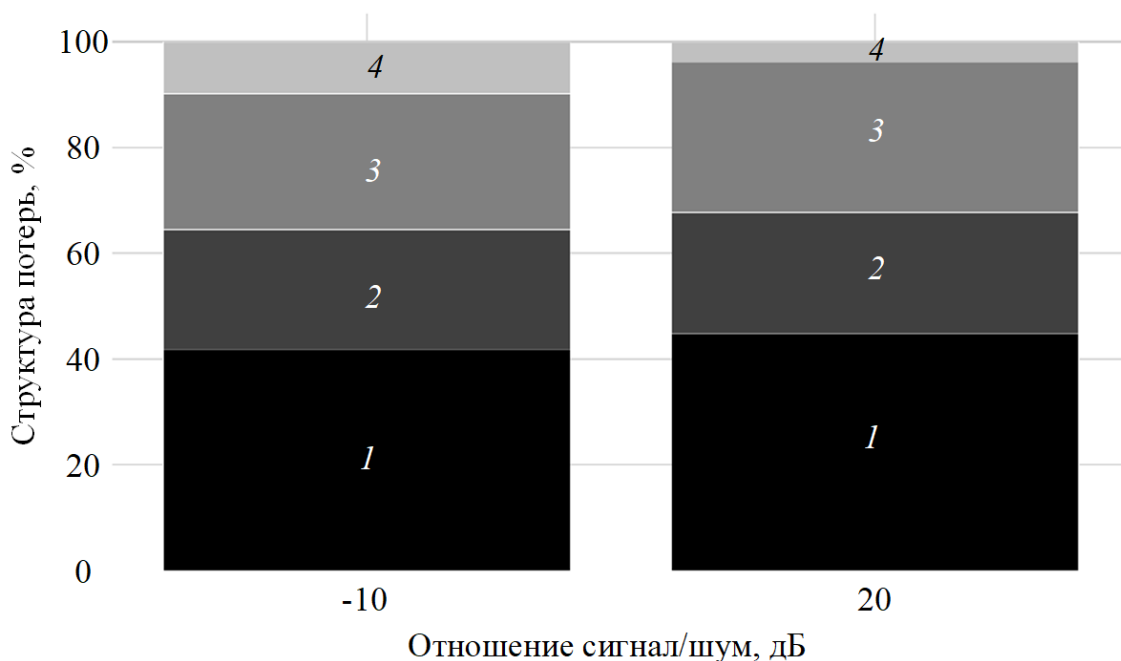


Рис. 7. Структура распределения потерь компонент многолучевого кластера (SNR –10 и 20 дБ, тестирование на RT-данных): 1 – ошибка классификатора (FN); 2 – потеря по временному разрешению; 3 – потеря по маскированию; 4 – прочие потери.

Для выявления механизмов потерь мощности на рис. 8 представлена кумулятивная структура мощности по категориям в зависимости от персонального SNR луча. Каждая линия показывает суммарную долю мощности всех компонент с персональным SNR ниже данного значения, отнесенную к соответствующей категории. На рис. 8 кривые 1-4 соответствуют категориям: 1 – успешно распознано (TP); 2 – ошибка классификатора (FN); 3 – потеря по временному разрешению; 4 – потеря по маскированию.

Ключевое наблюдение: линия маскирования лежит на оси абсцисс ($\approx 0,4\%$) – вклад этой категории в потерю мощности ничтожен. Прочие потери также имеют крайне низкий показатель мощности ($\approx 0,35-0,7\%$). Основные потери мощности приходятся на ошибки классификации (15-16 %) и потери по временному разрешению (7-8 %). Сопоставление с данными рис. 7 (где потери по маскированию составляют 26-28 % от суммарных потерь по числу лучей) показывает, что эти лучи несут менее 0,4 % мощности кластера – именно потому, что они маскируются боковыми лепестками сильнейшей компоненты.

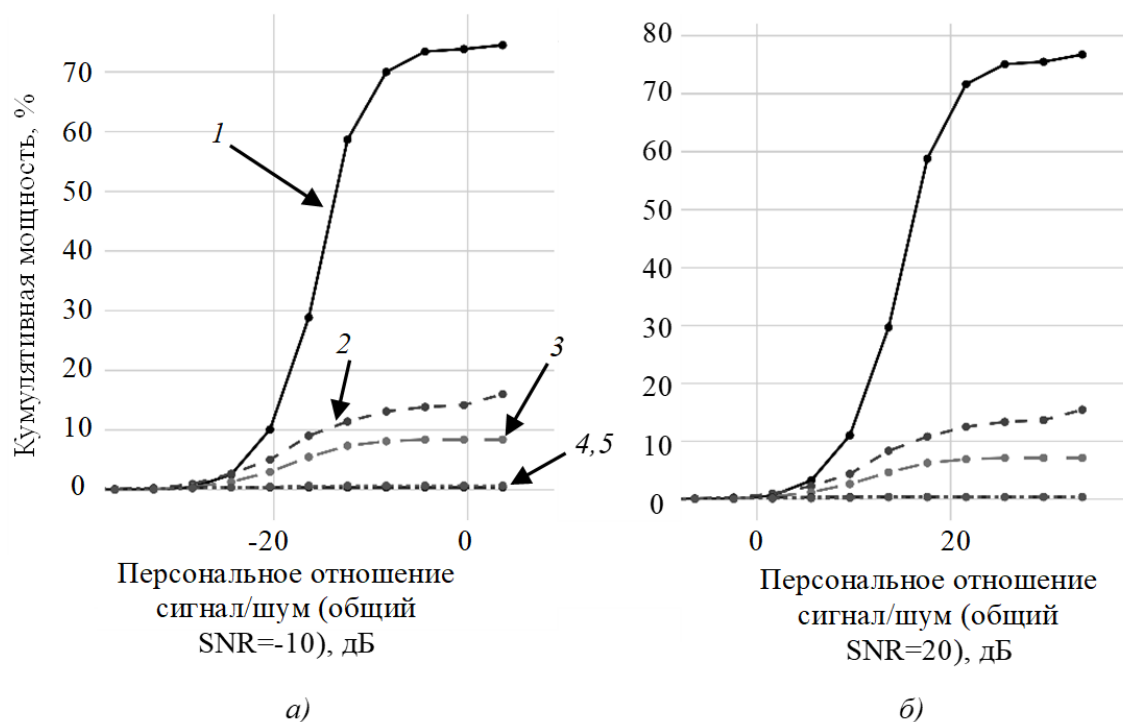


Рис. 8. Кумулятивная структура мощности по категориям (RT→RT):
 1 – успешно распознано (TP); 2 – ошибка классификатора (FN);
 3 – потеря по временному разрешению; 4 – потеря по маскированию;
 5 – прочие потери

Таблица 6. Кумулятивная мощность по категориям потерь (RT→RT, все сцены).

Категория	SNR луча = 5 дБ, SNR сигнала = 20 дБ	SNR луча = макс., SNR сигнала = 20 дБ	SNR луча = -25 дБ, SNR сигнала = -10 дБ	SNR луча = макс., SNR сигнала = -10 дБ
ТР (успешно), %	3,2	76,8	2,6	74,5
Ошибка классиф. (FN), %	2,2	15,4	2,7	16,0
Врем. разрешение, %	1,2	7,1	1,3	8,4
Маскирование, %	0,36	0,36	0,39	0,39
Прочие, %	0,14	0,34	0,38	0,69

Табл. 6 количественно подтверждает, что основные потери мощности связаны не с числом пропущенных лучей, а с их энергетической значимостью. При этом распределение вкладов по категориям остается практически неизменным при изменении SNR, что указывает на устойчивость выявленной структуры потерь.

4.3. Влияние геометрии среды распространения

На рис. 9 показана вероятность правильного распознавания компонент многолучевого кластера в зависимости от SNR для трех рассмотренных

городских сцен. На рис. 9 кривые соответствуют сценам: 1 – Флоренция; 2 – Сан-Франциско; 3 – Мюнхен. Повышение SNR с -10 дБ до 20 дБ дало незначительный прирост вероятности идентификации. Это указывает на то, что ошибки второго рода вызваны не столько аддитивным шумом, сколько маскированием слабых лучей и ограниченным временным разрешением корреляционной функции дальномерного кода. На рис. 10 представлена аналогичная зависимость для вероятности ложного обнаружения; кривые 1-3 соответствуют сценам: 1 – Сан-Франциско; 2 – Флоренция; 3 – Мюнхен. При $\text{SNR} = -10$ дБ вероятность FP достигала $2,1-3,4$ %, а при $\text{SNR} \geq 0$ дБ стабилизировалась на уровне $0,8-1,3$ %.

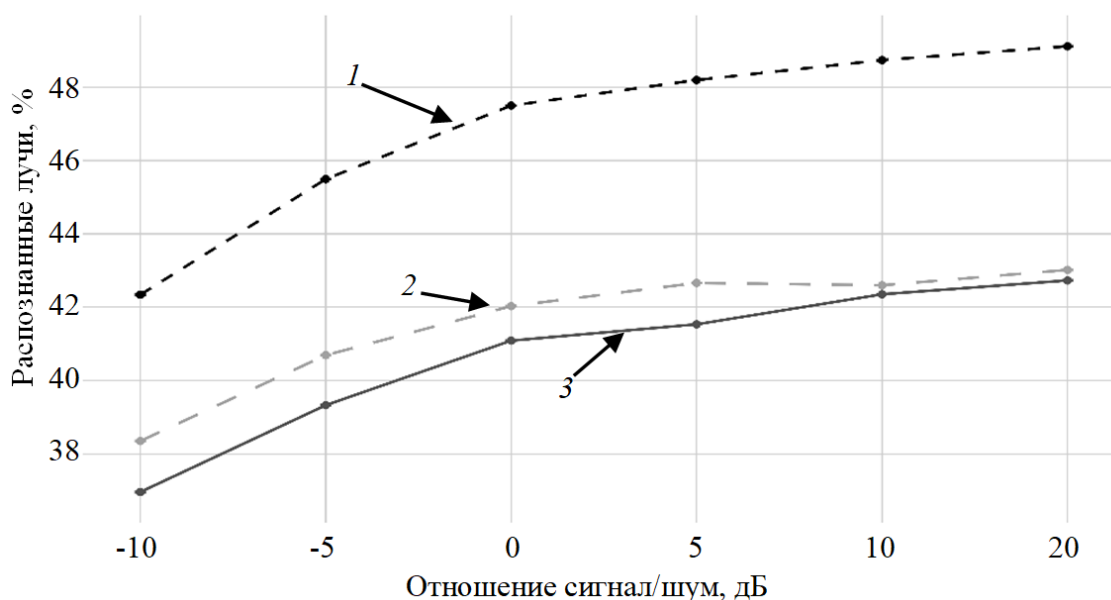


Рис. 9. Доля правильно распознанных компонент многолучевого кластера от SNR для различных RT-сцен (RT→RT): 1 – Флоренция; 2 – Сан-Франциско; 3 – Мюнхен.

Флоренция демонстрирует наибольшую вероятность распознавания ($42-49$ %), Сан-Франциско – промежуточную ($38-43$ %), Мюнхен – наименьшую ($37-43$ %). Мюнхен демонстрирует наименьший уровень FP при высоких SNR ($\approx 0,8$ %): при большом числе сильных реальных лучей классификатор реже принимает шумовой отклик за истинную компоненту.

Полученные значения FP при $\text{SNR} \geq 0$ дБ не превышают 1,5 %, что указывает на практическую применимость метода для задач формирования согласованных последовательностей.

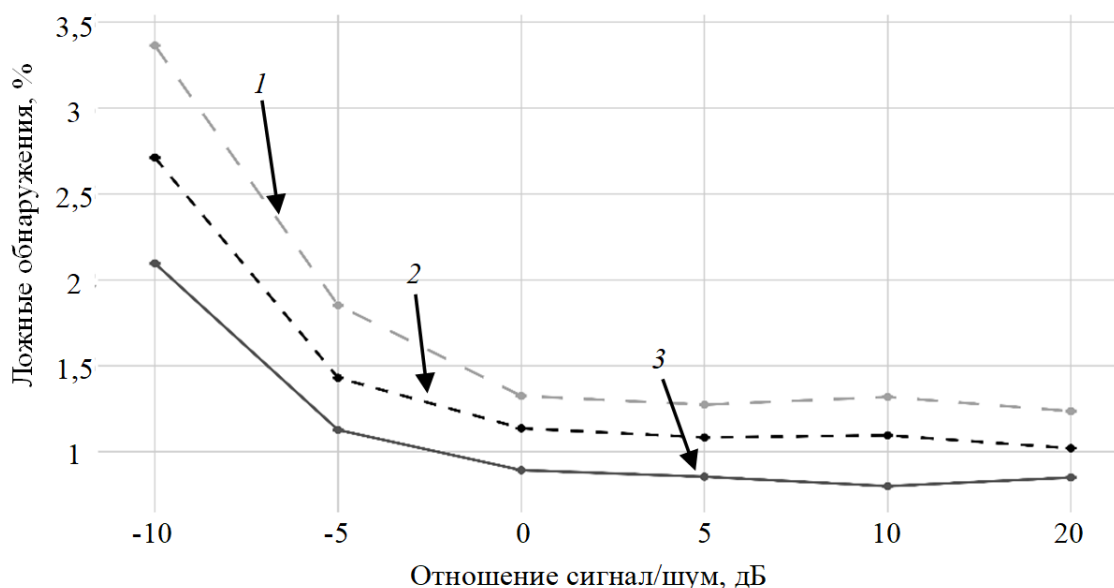


Рис. 10. Доля ложных обнаружений (FP) от SNR для различных RT-сцен (RT→RT): 1 – Сан-Франциско; 2 – Флоренция; 3 – Мюнхен.

Таблица 7. Сводные метрики по RT-сценам (агрегировано по всем SNR, RT→RT).

Сцена	Мощн., %	MAE, нс	Мощн. (N = 2), %
Флоренция	60,3	1,17	54,6
Мюнхен	73,0	1,20	62,5
Сан-Франциско	83,8	1,09	77,3

Геометрия застройки определяет энергетическую значимость лучей. В малоэтажной Флоренции распознается больше компонент (42-49 %), но они несут лишь 60 % мощности кластера. В высотном Сан-Франциско доля найденных лучей ниже (38-43 %), однако они аккумулируют 84 % энергии: алгоритм уверенно выделяет сильные доминирующие отражения, упуская слабые переотражения плотной среды.

Повышенный уровень ложных обнаружений (FP) в Сан-Франциско (рис. 10) обусловлен наличием мощного луча прямой видимости (LoS). При фиксированном общем SNR доминирующий LoS-луч забирает основной энергетический бюджет. Из-за этого эффективный SNR многолучевого кластера резко снижается,

отраженные компоненты опускаются к уровню шума, и классификатор чаще ошибочно принимает шумовые выбросы за истинные отклики.

Заключение

В работе проведено исследование эффективности идентификации парциальных компонент многолучевого кластера импульсной характеристики радиоканала по данным трассировки лучей для трех городских сцен различного типа в контексте задачи формирования согласованных случайных последовательностей (ССП).

Из полученных результатов следует, что для задач формирования СПП рекомендуется использовать не более 2-4 наиболее доминирующих лучей в канале. Это позволяет охватить порядка 66-69 % мощности рассеянного кластера при приемлемой вероятности ложного обнаружения $FP = 1,2-2,0$ %. Показано, что обучение классификатора на синтетических абстрактных данных TDL не обеспечивает требуемой надежности классификации (FP возрастает до 5-10 %), поэтому необходим учет особенностей реальной городской застройки.

За счет совместной корреляционной обработки и вейвлет-фильтрации сигнала удалось нивелировать эффект аддитивного шума, который не оказывает определяющего влияния на точность классификации: при экстремальном снижении SNR с 20 дБ до минус 10 дБ вероятность успешной идентификации снизилась всего на 5-7 %, а вероятность ложного обнаружения повысилась лишь на 1,3-2,1 %. Дальнейшее повышение производительности системы возможно за счет увеличения символьной скорости зондирования. Вместе с тем, наращивание длины дальномерного кода на практике строго ограничено временем когерентности радиоканала: превышение этого интервала приведет к изменению ИХ за время передачи и, как следствие, к разрушению корреляционных свойств принимаемого сигнала.

Установлено, что геометрия городской среды влияет преимущественно не на вероятность обнаружения лучей, а на распределение их энергии. В условиях исторической малоэтажной застройки (Флоренция) распознается

большее число компонент (42-49 %), однако они сильно рассеяны и описывают лишь около 60 % мощности. В то же время в условиях высотной застройки (Сан-Франциско) алгоритм выделяет меньше переотражений (38-43 %), но за счет уверенной идентификации сильных доминирующих лучей обеспечивает охват почти 84 % энергии многолучевого кластера.

Финансирование: Работа выполнена за счет средств Программы стратегического академического лидерства Казанского (Приволжского) федерального университета

Литература

1. Securing NextG networks with physical-layer key generation: A survey / Q. Xiao, J. Zhao, S. Feng, G. Li, A. Hu // Security and Safety. – 2024. – Vol. 3. – P. 2023021. <https://doi.org/10.1051/sands/2023021>
2. Physical Layer Secret Key Generation in Static Environments / N. Aldaghri, H. MahdaviFar // IEEE Transactions on Information Forensics and Security. – 2020. – Vol. 15. – P. 2692–2705. <https://doi.org/10.1109/TIFS.2020.2974621>
3. Key Generation From Wireless Channels: A Review / J. Zhang, T. Q. Duong, A. Marshall, R. Woods // IEEE Access. – 2016. – Vol. 4. – P. 614–626. <https://doi.org/10.1109/ACCESS.2016.2521718>
4. Physical layer key generation in wireless networks: challenges and opportunities / K. Zeng // IEEE Communications Magazine. – 2015. – Vol. 53, no. 6. – P. 33–39. <https://doi.org/10.1109/MCOM.2015.7120014>
5. Histogram-Based Estimation of CIR Entropy for Multipath Propagation Environments / A. V. Eponeshnikov, A. I. Sulimov // 2024 International Conference on Electrical Engineering and Photonics (EExPolytech). – IEEE, 2024. – P. 163–166. <https://doi.org/10.1109/EExPolytech62224.2024.10755705>
6. XGBoost: A Scalable Tree Boosting System / T. Chen, C. Guestrin // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – San Francisco, California, USA : ACM, 2016. – P. 785–794. <https://doi.org/10.1145/2939672.2939785>

7. 3GPP TR 38.901 Channel Model / Q. Zhu, C.-X. Wang, B. Hua, M. Kai, S. Jiang, M. Yao // Wiley 5G Ref. – 2021. – P. 1–35. <https://doi.org/10.1002/9781119471509.w5gref048>
8. Modeling the Wireless Propagation Channel: A Simulation Approach with MATLAB® / F. Pérez Fontán, P. Mariño Espiñeira. – 1st ed. – Wiley, 2008. <https://doi.org/10.1002/9780470751749>
9. 5G 3GPP-Like Channel Models for Outdoor Urban Microcellular and Macrocellular Environments / K. Haneda, J. Zhang, L. Tan, G. Liu, Y. Zheng, H. Asplund, J. Li, Y. Wang, D. Steer, C. Li, T. Balercia, S. Lee, Y. Kim, A. Ghosh, T. Thomas, T. Nakamura, Y. Kakishima, T. Imai, H. Papadopoulos, A. Ghosh // 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring). – IEEE, 2016. – P. 1–7. <https://doi.org/10.1109/VTCSpring.2016.7503971>
10. Sionna RT: Differentiable Ray Tracing for Radio Propagation Modeling / J. Hoydis, F. A. Aoudia, S. Cammerer, M. Nimier-David, N. Binder, G. Marcus, A. Keller // 2023 IEEE Globecom Workshops (GC Wkshps). – IEEE, 2023. – P. 317–321. <https://doi.org/10.1109/GCWkshps58843.2023.10465179>
11. TR 138 901 – V18.0.0 – 5G; Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 version 18.0.0 Release 18). – 3GPP, 2023.
12. Least-Squares-Based Iterative Multipath Super-Resolution Technique / W. Nam, S.-H. Kong // IEEE Transactions on Signal Processing. – 2013. – Vol. 61, no. 3. – P. 519–529. <https://doi.org/10.1109/TSP.2012.2225056>
13. A Superresolution Multipath Estimation Algorithm for DSSS Systems / F. Wang, W. Cui, J. Tian // IEEE Transactions on Aerospace and Electronic Systems. – 2023. – Vol. 59, no. 1. – P. 109–124. <https://doi.org/10.1109/TAES.2022.3188237>
14. A Survey of Artificial Intelligence Enabled Channel Estimation Methods: Recent Advance, Performance, and Outlook / B. Li, Q. Zheng, X. Tian, M. Yang, G. Gui, W. Jiang, H. Lei, J. Jiang, F. Shu, A. Elhanashi, S. Saponara // Artificial Intelligence Review. – 2025. – Vol. 58, no. 6. – P. 187. <https://doi.org/10.1007/s10462-025-11202-0>

15. Analysis of Machine Learning-Based NLOS Signal Identification Algorithm for UWB Indoor Localization Using CIR Waveform Features / K. Wang, C. Yang // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2024. – Vol. XLVIII-4-2024. – P. 705–710. <https://doi.org/10.5194/isprs-archives-XLVIII-4-2024-705-2024>
16. Fusing XGBoost and SHAP Models for Maritime Accident Prediction and Causality Interpretability Analysis / C. Zhang, X. Zou, C. Lin // Journal of Marine Science and Engineering. – 2022. – Vol. 10, no. 8. – P. 1154. <https://doi.org/10.3390/jmse10081154>
17. Local Predictability Analysis-Based Significant Multipath Component Identification for OFDM Systems / Y. Sui, H. Gao, Y. He, G. Jiang // IEEE Wireless Communications Letters. – 2024. – Vol. 13, no. 9. – P. 2482–2486. <https://doi.org/10.1109/LWC.2024.3420933>
18. A statistical basis for lognormal shadowing effects in multipath fading channels / A. J. Coulson, A. G. Williamson, R. G. Vaughan // IEEE Transactions on Communications. – 1998. – Vol. 46, no. 4. – P. 494–502. <https://doi.org/10.1109/26.664305>
19. Improved fading distribution for mobile radio / A. J. Coulson, A. G. Williamson, R. G. Vaughan // IEE Proceedings – Communications. – 1998. – Vol. 145, no. 3. – P. 197–202. <https://doi.org/10.1049/ip-com:19981991>
20. Radio Network Planning and Propagation Models for Urban and Indoor Wireless Communication Networks / W. J. Krzysztofik // Antennas and Wave Propagation / ed. P. Pinho. – InTech, 2018. <https://doi.org/10.5772/intechopen.75384>
21. A Deep Learning Approach to Position Estimation from Channel Impulse Responses / A. Niitsoo, T. Edelhäuser, E. Eberlein, N. Hadaschik, C. Mutschler // Sensors. – 2019. – Vol. 19, no. 5. – P. 1064. <https://doi.org/10.3390/s19051064>
22. Using the Physical Layer for Wireless Authentication in Time-Variant Channels / L. Xiao, L. Greenstein, N. Mandayam, W. Trappe // IEEE Transactions on Wireless Communications. – 2008. – Vol. 7, no. 7. – P. 2571–2579. <https://doi.org/10.1109/TWC.2008.070194>

23. Свойства кодов Голда и неоднозначность в РСА / Т. Е. Федорович, Ф. В. Сергеевич // Радиотехнические и телекоммуникационные системы. – 2023. – № 4. – С. 41–50. <https://doi.org/10.24412/2307-1304-2023-4-41-50>
24. Visushrink Pretreatment for Image Compression / W. Fourati, M. Bouhlef // International Journal of Computer Applications. – 2011. – Vol. 23, no. 6. – P. 1–5. <https://doi.org/10.5120/2893-3785>
25. Urban Radio Access Networks / S. Salous, T. Werthmann, G. Dahman, J. Flordelis, M. Peter, S. Hur, J. Park, D. Rose // Wiley 5G Ref. – 2022. – P. 17–69. <https://doi.org/10.1201/9781003337720-2>

Для цитирования:

Епонешников А.В., Сулимов А.И. Идентификация парциальных компонент импульсной характеристики радиоканала по данным трассировки лучей реальной городской застройки. // Журнал радиоэлектроники. – 2026. – №. 8. <https://doi.org/10.30898/1684-1719.2026.8.7>