

Двухступенчатый метод классификации активности человека в условиях отсутствия прямой видимости (NLoS) с использованием Wi-Fi CSI на базе ESP32

И. К. Калашников

*Санкт-Петербургский политехнический университет
Петра Великого*

kalashnikov.ik@edu.spbstu.ru

С. В. Завьялов

*Санкт-Петербургский политехнический университет
Петра Великого*

zavyalov_sv@spbstu.ru

Аннотация. В работе рассматривается задача бесконтактного распознавания активности человека с использованием Wi-Fi сигнала в условиях отсутствия прямой видимости. Система базируется на бюджетном микроконтроллере ESP32, что накладывает ограничения на качество сигнала и вычислительные ресурсы. Чтобы снизить влияние случайных сдвигов во временной области был реализован алгоритм линейной санитизации фазы. С целью оптимизации вычислений введен алгоритм выбора наиболее информативных поднесущих, который позволил сократить объем данных. Представлен алгоритм двухступенчатой классификации с использованием методов машинного обучения, который позволил снизить количество ложных срабатываний системы для статичных классов. Экспериментальная проверка показала, что предложенный метод обеспечивает точность распознавания активности 95.7% для четырех классов состояний.

Ключевые слова: классификация активности; ESP32; машинное обучение; Wi-Fi; оптимизация вычислений; информация о состоянии канала; распознавание активности человека; бесконтактные сенсоры

I. ВВЕДЕНИЕ

Распознавание активности человека становится все более популярным в таких сферах повседневной жизни, как «умные дома» [1] [2], мониторинг состояния здоровья человека [3] и в различных системах аутентификации пользователей. Нынешние системы мониторинга активности человека создают неудобства в повседневной жизни, поскольку требуют специальных сенсоров [4]. Системы, основанные на компьютерном зрении [5], работают только в условиях прямой видимости (Line-of-Sight) и из соображений приватности могут размещаться не везде.

С развитием беспроводных сетей появляется возможность использовать свойства радиосигнала (проницаемость, чувствительность к изменению среды) для решения вышеупомянутых проблем. Системы на базе Wi-Fi Sensing дают возможность оценивать активность человека даже в условиях отсутствия прямой видимости (NLoS), где традиционные методы неэффективны [6].

Целью данного исследования является разработка эффективного метода распознавания активности

человека в условиях NLoS на основе двухступенчатой классификации статистических признаков Wi-Fi CSI (Channel State Information, Информация о Состоянии Канала) на базе программируемого микроконтроллера ESP32 с внешней антенной 2.4 ГГц.

Для решения поставленной цели были решены следующие задачи:

- Теоретическое моделирование: Создана модель распространения Wi-Fi сигнала в среде MATLAB для анализа искажений амплитуды и фазы при прохождении радиоволн через движущийся объект (человека) в условиях многолучевости.
- Сбор данных: Сформирован набор данных CSI для различных видов активности (Empty, Sit, Stand, Walk).
- Предобработка: Реализован алгоритм фазовой санитизации (Phase Sanitization) для устранения аппаратных ошибок синхронизации между приемником (RX) и передатчиком (TX).
- Анализ признаков: Исследован набор статистических признаков во временной области. Выявлен наиболее устойчивый к затуханию через стены набор.
- Разработка алгоритма: Разработан и реализован алгоритм двухступенчатой классификации на основе машинного обучения (Machine Learning).
- Сравнение и валидация: Приведен сравнительный анализ эффективности ML-алгоритмов, который в будущем будет использован для оптимизации вычислений на периферийных устройствах сети (Edge Computing).

II. ПОСТАНОВКА ЭКСПЕРИМЕНТА

A. Конфигурация помещения

Эксперимент проводился в жилой квартире многоквартирного дома для того, чтобы достигнуть максимального уровня реалистичности сценария. Обстановка включает в себя крупногабаритную мебель, формируя сложную многолучевую среду.

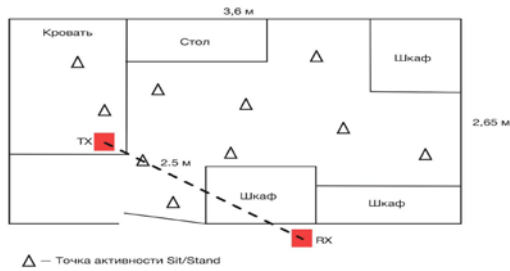


Рис. 1. Схема помещения

Система развернута в конфигурации «Сквозь стену» (Through-the-Wall).

- Передатчик (TX): Wi-Fi роутер, расположен внутри помещения
- Приемник (RX): Модуль ESP32, расположен снаружи, за бетонной стеной толщиной 15 см.

Расстояние между приемником и передатчиком составляет 2.5 метра.

В. Сбор данных

Сбор данных производился с помощью инструмента ESP32 CSI Toolkit [7]. Для исключения переобучения классификатора сбор данных для классов Sit и Stand производился в распределенных точках. Отношение классов примерно 1:1:1:0.7 (Empty:Sit:Stand:Walk). Данные записывались сессиями длиной 3 минуты с частотой дискретизации около 16 Гц. Несущая частота 2.4 ГГц (Канал 6, ширина полосы 20 МГц). Использовался микроконтроллер ESP32U с подключенной внешней антенной.

С. Классы активности

В выборку вошли 4 класса активности, охватывающие основные модели поведения человека в помещении:

- Empty (Пустая комната): В помещении отсутствует человек. Вне помещения происходит произвольная имитация различной деятельности человека.
- Sit (Человек сидит): Испытуемый сидит на стуле в естественной позе. Допускались микродвижения (Поворот головы, использование смартфона), имитирующие реальную деятельность.
- Stand (Человек стоит): Испытуемый стоит относительно неподвижно (Допускаются микродвижения).
- Walk (Произвольное перемещение): Испытуемый шагом перемещается по комнате.

Общая длительность записи данных составила около двух часов.

III. ОБРАБОТКА ДАННЫХ И ОБУЧЕНИЕ КЛАССИФИКАТОРА

А. Предобработка данных

Для оптимизации процесса сбора и уменьшения количества паразитных параметров были выделены не коррелирующие поднесущие (Subcarriers) OFDM сигнала, что позволило в значительной степени уменьшить объем полученных данных [8].

Первичная обработка включает в себя нахождение амплитуды ($|H(f, t)|$) и фазы ($\angle H(f, t)$) каждой

поднесущей. Поскольку оборудование не имеет синхронизации по времени необработанные данные CSI сильно искажены случайным смещением во временной области, следовательно, фаза нуждается в санитизации (Phase Sanitization) [9]. Для устранения фазовых искажений применялся алгоритм линейной санитизации (Linear Phase Sanitization) [9].

Для амплитуды и очищенной фазы применяется медианный фильтр (Hampel filter) [10] и фильтр низких частот Баттерворта (Low-pass Butterworth filter) с частотой среза 4 Гц. Это позволяет избавиться от случайных выбросов и аппаратного высокочастотного шума.

После очистки непрерывный поток данных CSI сегментируется с помощью техники скользящего окна (Sliding Window). Размер окна был выбран 150 пакетов (приблизительно 9 секунд). Этого достаточно, чтобы зарегистрировать паттерн поведения человека. Для увеличения объема обучающей выборки использовалось перекрытие окон (Overlap) в 70%.

В. Извлечение признаков (Feature Extraction)

Для каждой поднесущей во временном окне вычисляется набор из четырех статистических характеристик:

- Среднее значение амплитуды (Mean Amplitude):

$$\mu_A = \frac{1}{K} \sum_{t=1}^K |H(f, t)|$$

Характеризует общий уровень энергии сигнала. В конфигурации NLoS является ключевым для детекции присутствия человека в помещении.

- Стандартное отклонение амплитуды (Std Amplitude):

$$\sigma_A = \sqrt{\frac{1}{K} \sum_{t=1}^K (|H(f, t)| - \mu_A)^2}$$

Отражает меру изменчивости сигнала. Движения человека вызывают флуктуации, значение метрики увеличивается.

- Размах амплитуды (Amplitude Range):

$$R_A = \max(|H(f, t)|) - \min(|H(f, t)|)$$

Характеризует динамический диапазон значений. Наиболее эффективен для распознавания интенсивных действий.

- Дисперсия фазы (Phase Variance):

$$\sigma_\phi^2 = \frac{1}{K} \sum_{t=1}^K (\phi(f, t) - \mu_\phi)^2$$

После санитизации фаза становится чувствительно даже к самым незначительным смещениям (дыхание, микродвижения). Наиболее эффективная для классификации статичных поз.

Таким образом, для каждого окна формируется вектор признаков (Feature Vector), имеющий размерность $D = N_{\text{признаков}} \times N_{\text{поднесущих}}$:

$$X_i = [\mu_{A,1}, \dots, \mu_{A,K}, \sigma_{A,1}, \dots, \sigma_{A,K}, R_{A,1}, \dots, R_{A,K}, \sigma_{\phi,1}^2, \dots, \sigma_{\phi,1}^2]$$

Итоговая обучающая выборка представляет собой матрицу M_i размером $N_{\text{окон}} \times D$.

С. Архитектура классификатора

Для решения задачи распознавания активности в условиях NLoS была разработана двухступенчатая модель классификации. Чтобы избежать ошибок ложного обнаружения, которые являются критичными для систем безопасности и мониторинга, было принято решение отойти от традиционного метода (один классификатор на все классы).

В качестве базового алгоритма машинного обучения был выбран Случайный Лес (Random forest). Этот алгоритм устойчив к шуму в данных [11] [12], имеет низкую вычислительную сложность на стадии принятия решения [11] и позволяет моделировать сложные нелинейные зависимости [11]. Эффективность этого классификатора была подтверждена эмпирически.

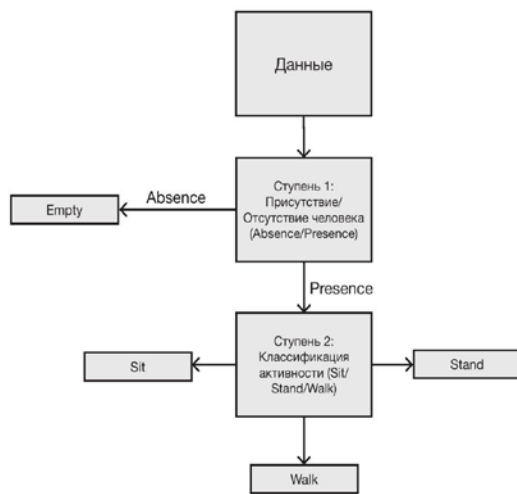


Рис. 2. Схема двухступенчатого алгоритма классификации.

На первом этапе классификатор фокусируется на поиске глобальных изменений в канале. Как показал анализ признаков, ключевым является средняя амплитуда, которая позволяет детектировать эффект затенения сигнала.

Поскольку наличие человека подтверждено, на втором этапе модель фокусируется только на динамике движений. Для различия статичных состояний возрастает вклад дисперсии фазы, так как эти паттерны слабо различимы из-за схожести динамики и микродвижений.

IV. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

А. Сравнительный анализ алгоритмов

Для оценки эффективности предложенного метода проведено сравнение методов машинного обучения.

ТАБЛИЦА I. СРАВНЕНИЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ.

Метод машинного обучения	Итоговая точность классификации	Время обучения
Random Forest	95.70%	~0.25 с
XGBoost	92.47%	~0.42 с
SVM	70.75%	~0.02 с
Logistic Regression	67.53%	~2.66 с
k-NN	63,87%	~0.01 с

Из таблицы видно, что алгоритмы на основе деревьев решений (Random Forest, XGBoost) показывают наилучшую точность (>92%) при этом значительно превосходя линейные и метрические. Это указывает на сложную нелинейную структуру границ классов в пространстве статистических признаков CSI. Как и ожидалось, Random Forest показал наилучший результат среди представленных методов, имея значительное преимущество по точности классификации.

В. Анализ точности классификации

Детальный анализ представлен в виде матрицы ошибок (Confusion Matrix).

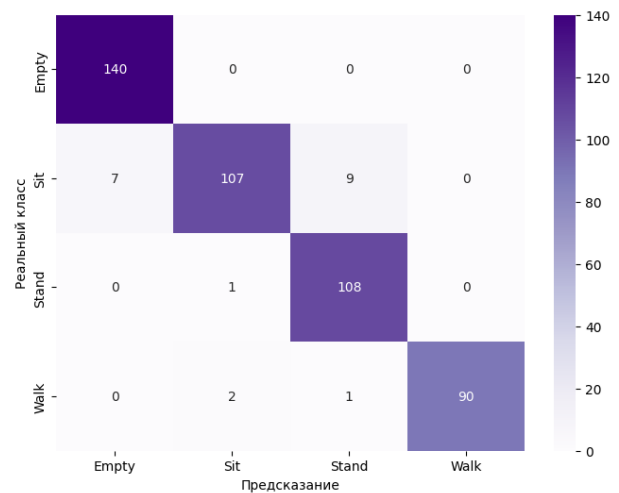


Рис. 3. Матрица ошибок классификации

Основная часть ошибок приходится на классы Empty/Sit и Stand/Sit. В условиях NLoS паттерны статичных классов могут быть схожи из-за похожих микродвижений и динамики, особенно при отдалении от зоны Френеля.

ТАБЛИЦА II. ОЦЕНОЧНЫЕ МЕТРИКИ КЛАССИФИКАЦИИ

F1-score	95.73%
Recall	95.71%
Precision	96.01%
AUC-ROC	99.21%

В табл. II приведены итоговые показатели эффективности. Значение F1-score на уровне 95.73% подтверждает баланс между точностью и полнотой обнаружения. Глобальная метрика AUC-ROC говорит о высокой статистической значимости результатов и устойчивости модели к вариативности входных данных.

С. Интерпретация признаков

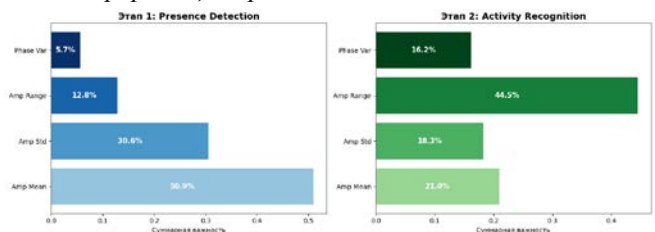


Рис. 4. Важность признаков на каждом этапе

На этапе детекции присутствия (Presence Detection) доминирующим признаком является Средняя Амплитуда (Amp Mean). Это подтверждает, что эффект затенения

имеет наибольшую значимость при детекции присутствия в условиях NLoS.

На этапе классификации активности (Activity Recognition) возрастает роль Размаха Амплитуды (Amp Range) и Дисперсии Фазы (Phase Var). Это доказывает, что для различения типа движения важна информация о фазовых сдвигах, которую удалось извлечь благодаря процедуре санитизации.

D. Визуализация пространства признаков

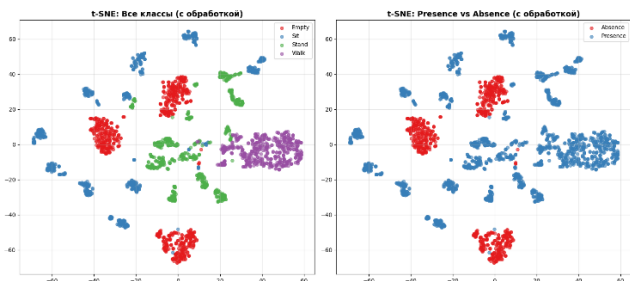


Рис. 5. Пространство признаков после обработки

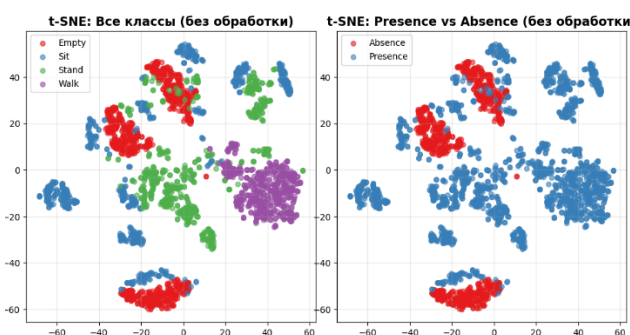


Рис. 6. Пространство признаков до обработки

Визуализация тестовой и тренировочной выборки с помощью t-SNE показывает эффективность выбранных признаков в условиях NLoS. Несмотря на высокую вариативность данных, перекрытия между классами на рис. 5 минимальны. Это говорит о том, что используемые признаки обеспечивают высокую обобщающую способность модели.

На рис. 5 класс Empty образует плотный изолированный кластер, что говорит о чувствительности системы к появлению человека. Классы Sit и Stand фрагментируются на несколько подгрупп, соответствующие разным точкам сбора данных в помещении. Такое распределение классов показывает невозможность проведения единой гиперплоскости для разделения кластеров одного класса от других, что в свою очередь доказывает неэффективность линейных моделей (Logistic Regression, SVM).

На необработанных данных (рис. 6) наблюдается сильное перекрытие классов. Без обработки случайные аппаратные шумы в значительной степени влияют на сигнал и кластеры становятся неразделимы. Применение предложенных алгоритмов значительно меняет топологию пространства признаков: наблюдается четкая кластеризация, границы становятся строгими, межклассовое расстояние возрастает.

V. ЗАКЛЮЧЕНИЕ

В работе представлен и экспериментально апробирован метод бесконтактного распознавания активности человека сквозь препятствия с использованием бюджетного микроконтроллера ESP32.

- Реализован метод обработки данных, включающий в себя систему фильтров и алгоритм фазовой санитизации, позволивший компенсировать низкое качество сигнала и ошибки синхронизации оборудования.
- Установлены наиболее информативные статистические признаки CSI, позволяющие четко разделить классы. Визуализация t-SNE подтвердила, что выбранный набор признаков дает высокую разделяемость классов.
- Предложенная двухступенчатая структура классификации на базе алгоритма Random Forest превзошла ожидаемый результат на ~10%. Итоговый результат: 95.70% точности для четырех классов состояний.
- Доказана возможность реализации пассивного мониторинга на устройствах с ограниченной вычислительной мощностью.

СПИСОК ЛИТЕРАТУРЫ

- [1] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in Proc. IEEE INFOCOM, Apr./May 2015, pp. 1472–1480.
- [2] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in Proc. ACM MobiCom, 2013, pp. 27–38.
- [3] C. Han, K. Wu, Y. Wang, and L. M. Ni, "WiFall: Device-free fall detection by wireless networks," in Proc. IEEE INFOCOM, Apr./May 2014, pp. 271–279.
- [4] C. Duan, X. Rao, L. Yang, and Y. Liu, "Fusing RFID and computer vision for fine-grained object tracking," in Proc. IEEE INFOCOM, May 2017, pp. 1–9.
- [5] Kalaiarasi, P., Kinnera Naveen, M. Venkata Arun Kumar, Varikuti Adhitya, and K. Lavanya. "A Real-Time Vision-based Fall Detection System using YOLOv10 and Deep Learning for Ensuring Elderly and Patient Safety." In 2025 6th International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 897-901. IEEE, 2025.
- [6] X. Fu, C. Wang and S. Li, "Wi-SensiNet: Through-Wall Human Activity Recognition Based on WiFi Sensing," 2024 5th International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), Wenzhou, China, 2024, pp. 238-241, doi: 10.1109/ICBASE63199.2024.10762264.
- [7] Hernandez, Steven M. "ESP32 CSI Toolkit." GitHub, 2021, github.com/StevenMHernandez/ESP32-CSI-Tool.
- [8] W. Shi et al., "CIT-HAR: A High Accuracy and Lightweight Human Activity Recognition System Using CSI Heatmaps and a Hybrid Transformer Network," in IEEE Transactions on Instrumentation and Measurement, vol. 74, pp. 1-17, 2025, Art no. 2541317, doi: 10.1109/TIM.2025.3591868.
- [9] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. SpotFi: Decimeter Level Localization Using WiFi. SIGCOMM Comput. Commun. Rev. 45, 4 (October 2015), 269–282.
- [10] Ali, H.A.H., Seytnazarov, S. Human Walking Direction Detection Using Wireless Signals, Machine and Deep Learning Algorithms. Sensors 2023, 23, 9726.
- [11] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001).
- [12] Wager et al., "A Degrees of Freedom Explanation for Random Forest Success", JMLR 2020