

## **ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ ДЛЯ РАСПОЗНАВАНИЯ АКТИВНОСТИ И УЛУЧШЕНИЯ КАЧЕСТВА ПОЗИЦИОНИРОВАНИЯ В ПОМЕЩЕНИИ**



***П. А. Шлык***

*БГУИР, г. Минск, Республика Беларусь  
star-war-sandakan@mail.ru*

***П.А. Шлык***

*Окончил Белорусский государственный университет информатики и радиоэлектронники.*

**Аннотация.** Исследуются теоретические предпосылки применения методов машинного обучения к задачам распознавания типа активности человека (Human Activity Recognition, HAR) и повышения качества позиционирования в помещении в условиях гетерогенного аппаратного обеспечения: различных

конфигураций сенсоров и характеристик среды. Рассматриваются теоретические основания для интеграции HAR-классификатора с байесовскими фильтрами, для применения нейронных сетей для сглаживания шума среды RSSI-сигнала, а также для обеспечения межустройственной переносимости моделей посредством алгоритмической калибровки и методов адаптации домена. Сформулированы гипотезы, подлежащие экспериментальной проверке.

**Ключевые слова:** HAR, машинное обучение, позиционирование

**Введение.** Классические аналитические методы позиционирования (трилатерация с использованием фильтра Калмана) не используют весь объём доступной информации: данные инерциальных датчиков (акселерометр, гироскоп), контекст активности, исторические траектории. Методы машинного обучения потенциально позволяют извлечь дополнительную ценность из потока сенсорных данных.

Существующие коммерческие решения (IndoorAtlas, Navigine) используют элементы машинного обучения, однако построение комплексной системы, объединяющей контекстно-зависимую навигацию, робастную обработку сигналов и методы межустройственного переноса моделей, остаётся малоисследованной областью.

**HAR: распознавание типа активности.** Задача распознавания типа активности (HAR) в данном контексте формулируется как классификация многомерных временных рядов. На вход модели подается скользящее окно сенсорных измерений  $W = \{s_i\}_{i=t-T}^t$  на интервале длительностью  $T$ , где каждый вектор  $s_i = ax, ay, az, x, y, zT$  содержит показания акселерометра и гироскопа. В качестве рабочего диапазона предлагается  $T_{2,5}$  с, что соответствует характерным временным масштабам циклов базовых двигательных активностей. Целью является сопоставление данного окна с одним из дискретных классов физической активности  $u \in \{\text{стоит, идёт, бежит, сидит, поднимается, спускается}\}$ .

**Рассматриваемые архитектуры.** В рамках исследования для классификации временных рядов рассматриваются две базовые нейросетевые архитектуры, выбор которых обусловлен доступными датчиками:

**Архитектура 1 (Многослойная рекуррентная сеть LSTM):** Ориентирована на устройства с полным набором инерциальных датчиков (акселерометр и гироскоп). Структурно модель представляет собой каскад рекуррентных слоев с применением методов регуляризации. Выбор данной архитектуры основывается на способности сетей с долгой краткосрочной памятью эффективно анализировать многомерные временные ряды, фиксируя характерные закономерности движения с учетом пространственной ориентации устройства, которую обеспечивает гироскоп [12].

**Архитектура 2 (Гибридная нейронная сеть 1D-CNN-LSTM) [1]:** Исследуется преимущественно для устройств с ограниченным аппаратным обеспечением (только трехосевой акселерометр), где данные подвержены высокому уровню позиционного шума из-за неизвестной ориентации устройства. Архитектура включает одномерный сверточный слой с последующей подвыборкой (pooling) и рекуррентным блоком. Теоретическое обоснование гибридного подхода заключается в том, что сверточный слой выступает в роли робастного фильтра, автоматически извлекая инвариантные к шуму локальные признаки (форму сигнала), тем самым снижая размерность и сложность задачи для последующего анализа временной динамики слоями LSTM.

**Проблема гетерогенности устройств.** Фундаментальной проблемой использования единой обученной модели на нескольких устройствах являются различия в самих датчиках, ориентации датчиков на печатной плате, частотах дискретизации (20-200 Гц), динамическом диапазоне и уровне шума. В литературе зафиксировано, что модели, обученные на данных одного устройства, демонстрируют значительное падение точности на других [7].

Существенная часть указанных различий может быть устранена алгоритмически, без привлечения методов машинного обучения. Предлагается калибровочный подход: модель обучается на данных в некоторой канонической системе координат, а при развертывании

на конкретном устройстве применяется матрица аффинного преобразования, приводящая показания датчиков к каноническому виду. Параметры преобразования могут быть определены на основании спецификаций производителя (ориентация осей, масштабный коэффициент). Аналогично, различия в частотах дискретизации устраняются алгоритмической передискретизацией к единой частоте. Ограничение данного подхода - необходимость знания аппаратной спецификации целевого устройства, что, выполнимо для большинства серийных моделей.

Для остаточных различий, не поддающихся алгоритмической компенсации (индивидуальные шумовые характеристики, нелинейности датчиков), рассматриваются методы адаптации домена. В частности, подходы по выравниванию распределений признаков между устройствами и частотно-ориентированная адаптация спектральных характеристик [7, 8]. Вопрос о том, в какой мере алгоритмическая калибровка снижает потребность в ресурсоёмких методах адаптации домена, подлежит экспериментальной проверке.

**Гипотеза интеграции HAR с фильтром позиционирования.** Результат HAR-классификации может быть использован для адаптивного переключения модели движения в байесовском фильтре. Теоретическим обоснованием данного подхода служит тот факт, что при разных типах активности характеристики движения качественно различаются:

Таблица 1. Модель движения для определённого типа активности

Активность	Предполагаемая модель движения
Стоит / сидит	Модель покоя, минимальный шум процесса
Идёт	Модель константной скорости (CV) [10]
Бежит	Модель константного ускорения (CA) [10]
Подъём по лестнице	Расширение вектора состояния z-компонентой

Данный подход концептуально близок к структуре фильтра с взаимодействующими множественными моделями (Interacting Multiple Model, IMM), в котором переключение между моделями движения управляется внешним HAR-классификатором [13]. Гипотеза состоит в том, что такая интеграция снижает ошибку позиционирования при смене режима движения по сравнению с фильтром с фиксированной моделью. Следует отметить, что использование скользящего окна длительностью  $T$  вносит соответствующую задержку в определение типа активности, что может влиять на точность фильтра при резких сменах режима движения.

**Устранение влияния среды на сигнал.** RSSI-сигнал в помещениях подвержен систематическим искажениям ( $\approx 4-8$  дБ): многолучевое распространение, затухание при прохождении через препятствия, интерференция [2]. Классические методы фильтрации (медианный фильтр, скользящее среднее, фильтр Калмана) способны устранять высокочастотный шум, однако они не моделируют сложную пространственную структуру помех, специфичную для геометрии конкретного помещения.

**Нейросетевое моделирование среды.** В качестве основного подхода рассматривается использование шумоподавляющего автоэнкодера (1D-Denoising Autoencoder) для моделирования поведения сигнала: [14]

- Вход: окно шумных RSSI-измерений  $rt-w, \dots, rt$ .

- Выход: оценка «чистого» уровня RSSI.

- Функция потерь: минимизация среднеквадратичной ошибки (MSE) между предсказанием и эталонным расстоянием.

Задача архитектуры заключается не в прямой низкочастотной фильтрации, а в построении неявной нелинейной модели искажений сигнала. Сеть, обучаясь на наборе данных конкретной локации, извлекает закономерности (многолучевое замирание, мертвые зоны), характерные именно для этой среды. Гипотеза состоит в том, что обученная модель

среды способна реконструировать истинное расстояние с меньшей ошибкой, чем универсальные аналитические фильтры.

**Позиционирование на основе радиокарты.** Метод радиокарты (fingerprinting) предполагает сопоставление вектора значений RSSI от  $N$  точек доступа с двумерными координатами. Для аппроксимации этой функции рассматриваются глубокие нейронные сети (DNN) и алгоритмы градиентного бустинга [6]. Теоретическое преимущество данного подхода заключается в неявном учёте сложной геометрии помещения без построения аналитической модели распространения сигнала. Рассматривается гибридная схема, где метод радиокарты используется для грубой инициализации координат, а байесовский фильтр - для их сглаживания в реальном времени.

**Подходы к формированию датасета.** Качество и репрезентативность обучающей выборки во многом определяют итоговую точность и обобщающую способность моделей машинного обучения [9]. Использование существующих открытых наборов данных для разработки комплексной системы позиционирования и распознавания активности сопряжено с рядом ограничений. Во-первых, публичные HAR-датасеты содержат показания инерциальных датчиков, но не включают синхронных RSSI-измерений и истинных координат [3]. В свою очередь, базы данных для локализации в помещении не содержат высокочастотных IMU-потоков и разметки контекста активности [4].

Во-вторых, как было отмечено ранее, модель искажений сигнала жестко привязана к геометрии конкретного помещения, что делает нецелесообразным перенос пространственных RSSI-данных из сторонних источников. Кроме того, большинство открытых датасетов собраны с использованием ограниченного набора устройств при фиксированном положении смартфона (например, только на поясе или в кармане), что затрудняет обучение моделей, устойчивых к аппаратной гетерогенности и произвольной ориентации.

В связи с этим возникает необходимость формирования собственного набора данных, собранного *in situ*, а также применения алгоритмических методов для искусственного расширения его вариативности. Ниже рассматриваются подходы к решению этих задач.

**Синтетический датасет из системы локализации.** Использование существующей системы локализации (Indoor Positioning System, IPS) позволяет получать траектории движения с временными метками [5]. Дифференцирование этих координат позволяет оценить скорость пользователя, что может служить основой для автоматической разметки типа активности. Синхронизация таких пространственных данных с показаниями инерциальных датчиков потенциально обеспечивает формирование обогащенного датасета без необходимости ручной разметки.

**Контрастное обучение.** В рамках данного подхода из одного временного окна сигнала путем различных аугментаций формируются положительные пары. Процесс обучения кодировщика направлен на максимизацию согласованности представлений внутри таких пар и минимизацию сходства между независимыми окнами, что достигается за счет использования функции потерь InfoNCE [15].

Применимость контрастного обучения к задачам классификации временных рядов сенсорных данных показана в [16]. Ожидаемым результатом является формирование признаковых представлений, инвариантных к характеристикам конкретного устройства и его положению.

**Методы аугментации сенсорных данных.** Для искусственного обогащения обучающей выборки и повышения устойчивости моделей применяются методы аугментации временных рядов. Рассматриваемы подходы включают:

- добавление высокочастотного шума (jittering);
- нелинейное искажение временной шкалы (time warping);
- случайное пространственное вращение (rotation) векторов ускорения и угловой скорости;

- масштабирование амплитуды сигнала (magnitude scaling);
- перестановку временных сегментов в рамках одного окна (permutation).

Вопрос об оптимальном наборе методов аугментации для задачи классификации активности подлежит экспериментальному исследованию.

**Гипотезы для экспериментальной проверки. H1:** Интеграция HAR-классификатора с UKF (Unscented Kalman Filter), реализующим адаптивное переключение моделей движения (константной скорости, константного ускорения, покоя), снижает среднюю ошибку позиционирования по сравнению с UKF с фиксированной моделью [11].

**H2:** HAR-модель на основе модуля ускорения (Архитектура 2) обеспечивает приемлемую точность классификации базовых активностей на устройствах без гироскопа, достаточную для управления моделью движения в фильтре.

**H3:** 1D-Denoising Autoencoder для RSSI превосходит медианный фильтр и скользящее среднее по метрикам MAE/RMSE оценки расстояния в условиях многолучевого распространения.

**H4:** Алгоритмическая калибровка данных сенсоров посредством матричных преобразований обеспечивает переносимость HAR-модели между устройствами с приемлемой потерей точности, существенно снижая необходимость применения ресурсоёмких методов адаптации домена по сравнению с прямым переносом (direct transfer).

**H5:** Синтетическая разметка HAR-датасета через данные системы локализации (IPS) пригодна для обучения моделей при условии, что ошибка позиционирования IPS не превышает порогового значения, определяемого экспериментально (предположительно, порядка характерного размера шага - 0.5 - 1 м).

**H6:** Использование глубоких нейронных сетей для позиционирования на основе метода радиокарты обеспечивает более точную начальную оценку координат по сравнению с аналитическими моделями распространения сигнала в условиях сложной геометрии помещения.

**H7:** Предварительное контрастное обучение на неразмеченных данных сенсоров позволяет извлекать признаки, инвариантные к типу устройства, что повышает точность HAR-классификатора при переносе на новые устройства без необходимости сбора дополнительных размеченных данных.

**Заключение.** Основным результатом проведенного теоретического анализа является обоснование перехода от изолированных алгоритмов позиционирования и распознавания активности к единой контекстно-зависимой системе обработки.

Показано, что методы машинного обучения потенциально способны эффективно дополнять классические аналитические модели: нейросетевые архитектуры берут на себя извлечение робастных признаков из зашумленных данных сред с многолучевым распространением сигнала (моделирование RSSI) и гетерогенных сенсоров, в то время как байесовский фильтр обеспечивает математически строгую агрегацию этих данных в реальном времени.

Предложенный гибридный подход, а также методология алгоритмической калибровки и контрастного обучения для обеспечения межустройственной переносимости формируют теоретический фундамент для создания робастных систем локализации.

В качестве направления для дальнейших исследований рассматривается экспериментальная валидация выдвинутых гипотез в условиях реальных промышленных объектов с высоким уровнем зашумленности радиоэфира и непредсказуемой ориентацией пользовательских устройств.

#### Список литературы

- [1] Ordóñez F. J., Roggen D. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition // *Sensors*. 2016. Vol. 16(1). P. 115.
- [2] Zafari F. et al. A Survey of Indoor Localization Systems and Technologies // *IEEE Communications Surveys & Tutorials*. 2019. Vol. 21(3). Pp. 2568-2599.
- [3] Anguita D. et al. A Public Domain Dataset for Human Activity Recognition Using Smartphones // *ESANN*. 2013.
- [4] Torres-Sospedra J. et al. UJIndoorLoc: A New Multi-building and Multi-floor Database for WLAN Fingerprint-based Indoor Localization // *IPIN*. 2014.
- [5] Kang W., Han Y. SmartPDR: Smartphone-Based Pedestrian Dead Reckoning for Indoor Localization // *IEEE Sensors Journal*. 2015.
- [6] Hsieh C. et al. Deep Learning for Indoor Localization with Wi-Fi Fingerprinting // *IEEE Access*. 2019.
- [7] Ragab M. et al. AdaTime: A Benchmarking Suite for Domain Adaptation on Time Series Data // *ACM Trans. Knowledge Discovery from Data*. 2023.
- [8] He R. et al. Domain Adaptation for Time Series Under Feature and Label Shifts (Raincoat) // *ICML*. 2023.
- [9] Sambasivan N. et al. "Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI // *CHI '21*. 2021.
- [10] Li X. R., Jilkov V. P. Survey of maneuvering target tracking. Part I. Dynamic models // *IEEE Transactions on Aerospace and Electronic Systems*. 2003. Vol. 39(4). P. 1333-1364.
- [11] Wan E. A., Van Der Merwe R. The unscented Kalman filter for nonlinear estimation // *Proc. IEEE Symposium on Adaptive Systems for Signal Processing, Communications, and Control (ASSPCC)*. 2000. P. 153-158.
- [12] Hochreiter S., Schmidhuber J. Long Short-Term Memory // *Neural Computation*. 1997. Vol. 9(8). P. 1735 - 1780.
- [13] Blom H. A. P., Bar-Shalom Y. The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients // *IEEE Transactions on Automatic Control*. 1988. Vol. 33(8). P. 780 - 783.
- [14] Vincent P. et al. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion // *Journal of Machine Learning Research*. 2010. Vol. 11. P. 3371 - 3408.
- [15] van den Oord A., Li Y., Vinyals O. Representation Learning with Contrastive Predictive Coding // *arXiv preprint arXiv:1807.03748*. 2018.
- [16] Eldele E. et al. Time-Series Representation Learning via Temporal and Contextual Contrasting // *Proc. IJCAI*. 2021. P. 2352 - 2359.

#### Авторский вклад

Шлык Павел Алексеевич - формирование статьи.

## THEORETICAL FOUNDATIONS OF ML FOR ACTIVITY RECOGNITION

*P. A. Shlyk*

*BSUIR, Minsk, Belarus  
star-war-sandakan@mail.ru*

**Abstract.** The research examines the theoretical foundations for applying machine learning methods to human activity recognition tasks (Human Activity Recognition, HAR) and improving indoor positioning accuracy under heterogeneous hardware conditions, including various sensor configurations and environmental characteristics.

It addresses the theoretical basis for integrating a HAR classifier with Bayesian filters, using neural networks to smooth RSSI signal environmental noise, and ensuring inter-device model portability through algorithmic calibration and domain adaptation techniques.

Hypotheses have been formulated for experimental verification.

**Keywords:** HAR, ML, indoor positioning