RoomE and WatchBLoc: A BLE and Smartwatch IMU Algorithm and Dataset for Room-Level Localisation in Privacy-Preserving Environments

Ada Alevizaki(✉) and Niki Trigoni

Department of Computer Science, University of Oxford, Oxford, UK
{antigoni.alevizaki,niki.trigoni}@cs.ox.ac.uk

Abstract. The increasing at-home way of living, which became essential due to the COVID-19 pandemic, yields significant interest in analysing human behaviour at home. Estimating a person's room-level position can provide essential information to improve situation awareness in smart human-environment interactions. Such information is constrained by two significant challenges: the cost of required infrastructure, and privacy concerns for the monitored household. In this paper, we advocate that ambient bluetooth signals, from IoT devices around the house, and inertial data from a smartwatch can be leveraged to provide room-level tracking information without additional infrastructure. We contribute a comprehensive dataset that combines real-world BLE RSSI data and smartwatch IMU data from two environments, which we use to achieve room-level indoor localisation. We propose an unsupervised, probabilistic framework that combines the two sensor modalities, to achieve robustness against different device placements and effectively track the user around rooms of the house, and examine how different configurations of IoT devices can affect the performance. Over time, through transition-events and stay-events, the model learns to infer the user's room position, as well as a semantic map of the rooms of the environment. Performance has been evaluated on the collected dataset. Our proposed approach boosts the localisation accuracy from 67.77% on average in standard BLE RSSI localisation, to 81.53%.

Keywords: indoor localisation · semantic mapping · smartwatch · ambient IoT device · BLE · IMU · probabilistic graphical models · HMM · dataset

1 Introduction

In recent years, people spend a considerable amount of their day in their home. An increased tendency towards self-employment, as well as businesses experimenting with new productivity schemes, have given more flexibility to people to work from home, a trend that has largely culminated to a necessity due to the COVID-19

S. Longfei and P. Bodhi (Eds.): MobiQuitous 2022, LNICST 492, pp. 266–289, 2023.
https://doi.org/10.1007/978-3-031-34776-4_15

pandemic. At the same time, the advances in medicine and the increase in life expectancy have resulted in the phenomenon of the ageing population; retired adults normally have many more years to live, but are often faced with age-related diseases that might keep them at home as they become more severe.

The increasing at-home way of living presents a unique opportunity to analyse human behaviour at home to improve situation awareness for Cyber Physical Systems (CPS) that function based on human-environment interactions. For example, smart homes could benefit from daily patterns to learn tasks, such as preheating a room that is frequented at a usual time. Similarly, older adults could be remotely supervised, and intervention could be called for should an abnormal incident be detected.

By investigating such scenarios, it becomes evident that human-environment interactions at home are largely correlated with semantic locations around the house and the mobility patterns of the user around different rooms. As such, inferring the position of a human inside their house as well as transition patterns across rooms can significantly improve situation awareness.

However, human positioning at home is far from trivial. Installation of bespoke infrastructure can be cumbersome and cost-prohibitive. Privacy concerns can also be raised, not only regarding the person monitored, but also for others that might be sharing the space. It is therefore important to identify methods that can exploit existing infrastructure to infer usage of the space and effectively capture mobility patterns, in a privacy-preserving way. Despite the surge of research activity in recent years in the area of human indoor positioning, these constraints rule out many methods that are very effective in different context.

Infrared (IF) positioning systems, for instance, though capable of providing very accurate position estimates, require installation of extensive infrastructure, making them cost-prohibitive for the home environment case [24]. Pedestrian Dead Reckoning (PDR) is an alternative that could minimise costs significantly, as it only requires cheap inertial sensors (accelerometer, gyroscope, magnetometer) that can also easily be attached on the person to be monitored and are widely available in everyday devices, such as smartphones and smartwatches; however, it requires very accurate knowledge of one's initial position, and it can also exhibit large errors due to drift [11]. This issue is exacerbated in smartwatches where there could be interference from multiple other non walk-related hand motions [11]. Other radio frequency (RF) localisation techniques have successfully been used in indoor public spaces (e.g., shopping malls), that are equipped with large wireless sensor networks (WSN/WLAN) with numerous access points (APs). These methods usually employ Wi-Fi APs or bluetooth low energy (BLE) beacons [37] together with smartphone inertial measurement unit (IMU) sensors [32], to achieve an accuracy with an estimated $2 - 10$ metres error [32]; errors beyond 4 m however would probably correspond to a completely different room in a house. Furthermore, they usually require extensive fingerprinting or crowdsourcing [36], which is not an option at home. Lastly, even though mobile phones are carried very often by their owners wherever they go, it is not common for people to carry the device at all times inside their house. Elderly people in particular, are often unfamiliar and resilient to use such technology. Device-free

passive RF localisation, that does not require any sensors to be carried, could probably be effectively used to monitor people who live alone, but it will raise difficulties in more crowded houses, as there is no way to distinguish between movements of the occupants of the house (or visitors) if they are not carrying a tag, and requires careful training per house.

Furthermore, there are not many publicly available datasets that are suited to develop indoor positioning methods for the home scenario. In [18], the researchers collect geographical ground truth locations and BLE received signal strength indicators (RSSIs), but all data are collected in a single open space, with known device configurations and floorplan. Similarly, the UJIIndoorLoc-Mag database [27] provides semantic location data and smartphone IMU data, but is only constrained in snapshots of corridors, and not continuous movement across different rooms. The Miskolc IIS Hybrid IPS dataset [29] offers geographical ground truth location, BLE and Wi-Fi RSSIs, as well as magnetometer data from a smartphone, which constraints the viable positioning methods to those based on magnetic field (which usually require fingerprinting). The dataset collected with the Smart Home in a Box (SHIB) technology [17] is the most complete database, offering semantic location estimates and RSSIs from gateways and IMU data in a wrist wearable in a real home environment. Other datasets that offer smartwatch IMU recordings, e.g., [20,31], are not coupled with location information, as they are more suited to human activity recognition applications, than indoor localisation.

In this work we propose **RoomE**, a smartwatch-based method to achieve room-level localisation at home. Smartwatches are equipped with inertial sensors and Bluetooth radio, with which they can sense IoT devices in their vicinity. As these are seeing increasing use among the wider population, we assume this to be a practically infrastructure-free approach. We also introduce **WatchBLoc**, a new dataset of *smartwatch* IMU data and *BLE* RSSI, along with ground truth semantic *location* of the user, collected "in the wild" across two environments.

In summary, our main contributions are:

- We propose **RoomE**, a probabilistic algorithm that achieves semantic localisation and room identification, by sequentially fusing ambient location information with high-level mobility features derived from inertial data, no matter the IoT devices' placement in the environment.
- To our knowledge this paper introduces the first comprehensive dataset, **WatchBLoc**, comprising signals from BLE home devices and user-worn smartwatches allowing for the study of room-level localisation under a variety of conditions: users, IoT device layouts and home layouts.
- We investigate the challenges associated with the sensors we use in this study: i) how different placements of IoT devices around the house can affect the accuracy of room detection, and ii) the challenges associated with the use of motion data from smartwatches and the need for a simple yet robust approach to detecting candidate room change events.
- We evaluate our proposed approach on the new dataset across different floorplan layouts, IoT device placements and user motion patterns. We show that

our method mitigates the issue of high sensitivity of existing smartwatch room detection algorithms to the layout of ambient home devices. It increases the room detection accuracy compared to state of the art from 67.77% to 81.53% without making use of fingerprinting of the environment. It also provides the capability of inferring room transition probabilities without using floorplan information.

The remainder of the paper is organised as follows: Sect. 2 presents the newly introduced dataset and discusses challenges and lessons learnt in collecting as well as analysing the data for the indoor localisation scenario. Section 3 presents the architecture of our proposed indoor localisation system and illustrates how the different modules work together. Sections 4, 5 and 6 introduce the algorithmic details for each of the system's components, respectively, and evaluate their performance. Section 7 then discusses related state-of-the-art methods and the challenges each one faces. Finally, Sect. 8 concludes the paper and discusses limitations and future directions.

2 The Dataset: WatchBLoc

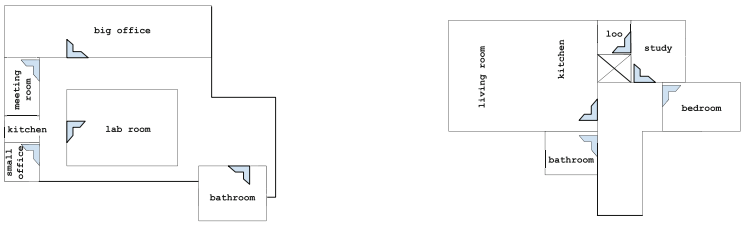
In this section we will present **WatchBLoc**, a new dataset that we collected to achieve indoor localisation for home environments using BLE RSSIs from ambient IoT signals and smartwatch IMUs.[1]

2.1 Dataset Description

We collected data from a number of participants recorded as they were freely performing their typical daily home activities. The data were recorded in two different environments: a *demo-home*, i.e., an office space that was set up to facilitate the experiment, and a *real-home*, i.e., a standard flat. The floorplans of the data collection environments are shown in Fig. 1. BLE beacons were placed in each room (including bathrooms, but excluding corridors), to simulate the existence of IoT devices that can normally be present in a house, like a BLE-connected TV or fridge. We assume that exactly one IoT device exists in each room of the house.

The *real-home* is an interesting case of a house to examine: it is a 2-bedroom/2-bathroom flat, with an open-plan kitchen-living room space. The bedroom denoted as `study` and the bathroom denoted as `loo` constitute an ensuite room. As such, even though there exist 6 different semantic locations (namely `kitchen`, `living room`, `bedroom`, `bathroom`, `study` and `loo`), these can be accounted as 5 or 4 distinct rooms: an open-plan `kitchen/living room`, a `bedroom`, a `bathroom`, and a `study` and `loo`, or a single `ensuite` space. We will discuss all 4 combinations in this work, denoted as *full*, *openplan*, *ensuite* and *merged*, respectively.

[1] The dataset is publicly available at: https://doi.org/10.5281/zenodo.7039554.



(a) *demo-home*. (b) *real-home*.

Fig. 1. Data collection environment layouts. The floorplans of the two environments where we collected data for our dataset are presented. The *demo-home* consists of $N = 6$ rooms, arranged across a single floor. The *real-home* is a 2-bedroom/2-bathroom flat, with an open-plan kitchen-living room space. There are 6 distinct semantic locations, that however can be accounted as $N = 4$, $N = 5$ or $N = 6$ rooms, depending on whether we assume the open-plan kitchen and living room to constitute a single *openplan* room, the ensuite study and loo as a single *ensuite* room, or both.

The experiment was performed by 11 participants split across the two environments, yielding a total of 20 continuous recordings, each containing data for three beacon configurations as detailed below. Details on each of the recordings are listed in Table 1.

Table 1. Description of **WatchBLoc** dataset. A total of 11 users performed the experiment across two different environments. `user1` performed the experiment in both the demo-home (for three distinct recording IDs, 1, 4 and 8) and the real-home (recIDs 13, 15, 17, 19). `user11` performed the experiment in the real-home (recIDs 14, 16, 18, 20), and the rest of the users in the demo-home, yielding one recID each. Four versions of semantic layout for the *real-home* are considered: each semantic location is considered a separate room (*real-home:full*); an ensuite bathroom and bedroom are abstracted to a single ensuite room location in the *real-home:ensuite* case. An open-plan kitchen and living room existing in a space are abstracted to a single semantic room, kitchen/living room in *real-home:openplan*. Finally, both the ensuite and open-plan abstractions are considered in the *real-home:merged* case.

recID	environment	configurations	# rooms
1-12	*demo-home*	*centre*, *doors*, *far*	$N = 6$
13-14	*real-home:full*	*centre*, *doors*, *far*	$N = 6$
15-16	*real-home:ensuite*	*centre*, *doors*, *far*	$N = 5$
17-18	*real-home:openplan*	*centre*, *doors*, *far*	$N = 5$
19-20	*real-home:merged*	*centre*, *doors*, *far*	$N = 4$

We examined three possible IoT device configurations: placing beacons at the centre of each room (*centre*), at the entrances of each room (*doors*), and at locations around each room such that their pairwise distances are maximised (*far*). Combinations of the above device configurations might exist in real life, e.g., there might be a smart thermostat by one room's door and a smart speaker in the centre of another; we only discuss the aforementioned three configurations in this paper, but the dataset includes BLE RSSIs from all the above $3N$ locations simultaneously, allowing researchers to examine other combinations of device configurations should they wish.

Each of the recordings consists of BLE and IMU data, as perceived by a Sony Smartwatch 3, as well as ground truth semantic location of the user as detailed below:

1. BLE RSSIs from any BLEs in the vicinity of the user, at a frequency of 0.2 Hz,
2. Inertial data (accelerometer, gyroscope, magnetometer) of the smartwatch, at a frequency of 100 Hz.
3. Ground truth location of which room the user was in, logged by the user by tapping the smartwatch screen when changing rooms.

2.2 Challenges

The collection of this dataset has brought to light a few challenges associated with seamless, privacy preserving localisation of people. We will discuss some of these here, but will mention further difficulties encountered in the following sections, as they appear in the dataset's analysis and algorithm's evaluation.

The BLE RSSI data from different houses may vary significantly, even for houses that are of similar layout. Wireless signal strength can be affected by many factors, such as the existence of line-of-sight (LOS) between a BLE device and the smartwatch, attenuation at different levels depending on the construction of the walls' material [8], etc. As such, it is not always guaranteed that we will be hearing a BLE RSSI as we would expect in a space; sometimes RSSI from an IoT device that exists in an adjacent room may be stronger than the RSSI from an IoT device located in the space the user is in. Or, we might have missing data (i.e., not detecting certain beacons at all) even if we're continuously recording data; this can be the result of severe signal attenuation owing to the house's layout (e.g., walls), but also of the user's and other occupants' movements around the house [7], that might be creating additional obstructions to the LOS between device and smartwatch.

A few other challenges arise from the choice of the smartwatch as the sole sensing modality. To ensure long battery life, smartwatches may go to battery saving mode, e.g., when their screen is not touched for a long time; this can result to large segments of missing data, which can lead to uncertain positioning estimates for the corresponding periods of time. At the same time, the wrist attachment of the smartwatch means that the watch is recording inertial data that encode information not only due to the user moving across or within rooms, but also when any other activity is performed, as hand usage is entailed in

most everyday activities [11]. This leads to a lot of additive "noise" and can significantly complicate distinguishing walking from other vigorous activities, such as brushing teeth or doing the dishes.

Ground truth collection whilst adhering to privacy concerns has also been a significant challenge during the design and data collection of this experiment. Initial attempts to data collection included ground truth collected in the form of hand-written notes from the user. This was impractical, as the exact timings of changing location and ground truth logs were almost impossible to align, and it was heavily relying on the participant to remember their exact course of actions. Camera-use was prohibited, as data collection included performing the experiment at participants' private houses and was thus considered a privacy breach by the Ethics Approval Committee. RF tags were also inappropriate to indicate ground truth location of the participant, as multiple people were using the data collection environments at the same time. Based on these issues, a logging application was developed and installed on the smartwatches, so that the participants could tap the appropriate semantic location on the smartwatch's screen every time they were moving to a new room. It is important to note that this approach can still have synchronisation issues, as it is almost impossible for the participant to tap their new location exactly at the moment of room-change. However, exploration of the data has shown that it is reasonable to assume that the lagging in the ground truth is no longer than approximately 5 seconds.

3 System Architecture: RoomE

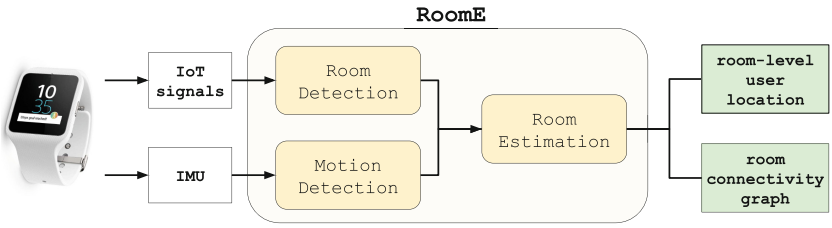


Fig. 2. System architecture

This section provides a high-level description of our system architecture and its three main algorithmic components, i.e., **Room Detection**, **Walking Detection** and **Room Estimation**, each of which will be detailed and evaluated in the following sections.

Our algorithm assumes a house with N rooms, each equipped with exactly one IoT device[2]. We also assume that we can infer a semantic label for the

[2] In practice this is not constraining in the presence of multiple IoT devices per room, as they can be aggregated into an e.g., 'max' or 'mean' BLE signal. Handling rooms with no BLE devices is more challenging and is an open topic for future work.

room that each IoT device is located in, through its identifiable name, e.g., an "IoT fridge" is located in "kitchen", etc. The inhabitant is in possession of a smartwatch, which can perceive these devices through their emitted signals. No information is available regarding the size of each room, their relative position in the house or the position of the IoT devices within the rooms.

The goal of our algorithm is to estimate the room-level location of the inhabitant as well as a semantic map of the environment, while they freely move around the house. This is achieved through **RoomE**'s three main components. These are depicted in Fig. 2 and outlined below:

1. **Room Detection**, giving initial estimates of the user's semantic location.
2. **Motion Detection**, to estimate potential room-transitions.
3. **Room Estimation**, to provide a definite, improved room-level user location and information on the house's layout.

The first two modules utilise the ambient IoT signals from the various IoT devices and the inertial data logged in the smartwatch, to derive initial estimates of the room-level semantic location and mobility of the user, respectively.

In the final module, we employ a probabilistic graphical model, which we train against the available data, to fuse the information extracted in the above modules in a probabilistic graph. Traversal of the graph in the inference phase provides refined room-level localisation for the user. The learnt model's parameters during the training phase are used to infer room-connectivity information, i.e., which rooms are accessible from others.

RoomE thus outputs a room-level location for the user at each timestamp, as well as an estimated semantic map of the house, in the form of a room-connectivity graph.

4 Room Detection

In this section we will discuss the details of the **Room Detection** module of our algorithm, and provide an analysis of the BLE RSSI data recorded in our dataset.

4.1 BLE Data Preprocessing

A basic preprocessing was used to prepare the BLE RSSI data from our collected dataset for the **Room Detection** module.

First, we segment each BLE RSSI signal in non-overlapping windows of length 5.12 s[3]. When multiple RSSI logs from the same device exist in a 5.12 s window, these are aggregated to a single RSSI value, the maximum among them. The timestamp t for the data in each window is thus now updated to be the end-time of the window. To deal with missing "no-signal" RSSI values, the BLE data

[3] The choice of the window size is related to the 0.2 Hz recording frequency of the BLE data, as well as an implementation detail in our approach for the **Motion Detection** module that eases time synchronisation of the two modalities.

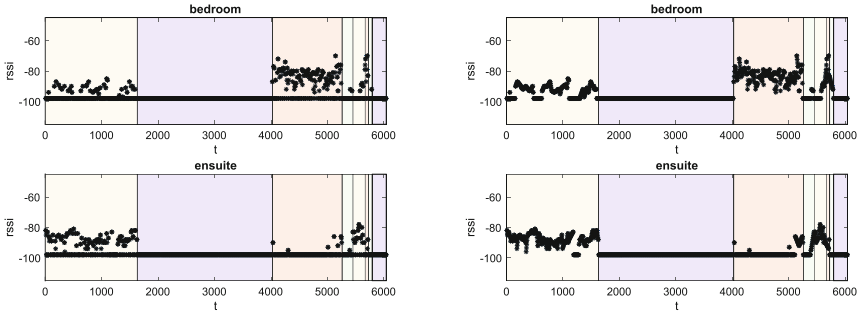
are linearly interpolated, following the 5.12 s max-aggregation. Figure 3 shows an example of the behaviour of RSSIs from two different rooms in *real-home*, as well as the effect of the preprocessing on the final BLE data, which are used by the rest of the **Room Detection** module.

4.2 Algorithmic Details

For the **Room Detection** module, we employ a *maxRSSI* approach on the preprocessed BLE data, thus assuming that at any point, the user is nearest to the IoT device that emits the strongest RSSI and thus in the corresponding room in the house where the IoT device is placed.

Note that other algorithms that are able to infer semantic location from ambient IoT data could very well be used instead of *maxRSSI*. Our choice of *maxRSSI*, is based on the constraints that the home environment poses: we do not have any information about the dimensions of the rooms or the IoT devices' placements around the house; this makes the use of e.g., fingerprinting based algorithms less suitable for our problem. Even if we did have such information, we would need to re-calibrate the model in each house, as the signal propagation would be affected by the specific conditions of each environment.

maxRSSI is a simple approach that does not require prior training or calibrating, as it is only concerned with the strongest RSSI signal at each time. It has been considered accurate enough to generate the ground-truth location of users in human-robot interactions [21], as well as to identify the interactions of users with objects that have BLE beacons attached [13]. It also resembles



(a) Raw BLE signals. (b) Preprocessed BLE signals.

Fig. 3. Example of raw and preprocessed BLE RSSIs across rooms. BLE RSSIs logged in two adjacent rooms, `bedroom` and `ensuite`, in the *real-home* environment of our dataset. The coloured areas denote the different ground truth room the user was in at each time, each colour corresponding to a separate room; pink and yellow in particular, correspond to ground truth bedroom and ensuite locations, respectively. We observe that in adjacent rooms, we can register RSSIs from multiple BLEs. Also, raw BLE RSSI recordings (left) can have missing data. Interpolation can help in smoothing intermittent no-signal phases (right).

the thought process behind log-linear path-loss radio propagation models, which are often employed to model the degrading effect of the distance between the transmitter and the receiver on the RSSI values:

$$r = r_0 - 10n \log_{10}(d),$$

where r is the RSSI at the distance d, r_0 is the received signal power of the receiver from a transmitter one meter away and n is the path loss exponent. As is obvious from the above, for $d_1 > d_2$ it holds that $r_1 < r_2$, i.e., the RSSI value decreases as the distance from the transmitter increases. As such, the maximum RSSI value will be obtain nearest to the transmitter.

As discussed in Sect. 4.1, the data passed to the *maxRSSI* algorithm are vectors of RSSIs, corresponding to the N RSSIs received from the N identified rooms in the house at each timestamp t (i.e., the end-time of the corresponding window):

$$vecRSSIs_t := [rssi_{r_1,t}, \ldots, rssi_{r_N,t}] .$$

The room estimate r_t can then be computed as the room where the strongest signal device is located. For windows when, even after the preprocessing, no signal was received, we cannot have an accurate estimate for the user's location with the *maxRSSI*. As such, we assign these windows to an "unknown room":

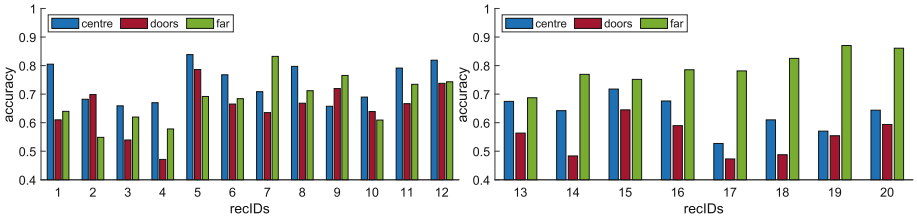
$$r_t = \begin{cases} \text{unknown room}, & \text{if no RSSI from beacons,} \\ r_j,\ j = \operatorname{argmax} vecRSSIs_t, & \text{otherwise.} \end{cases}$$

4.3 Evaluation

We evaluate the above *maxRSSI* approach, for the **Room Detection** module on the BLE RSSI data and the corresponding ground truth semantic location, as they were collected in our dataset.

Figure 4 demonstrates the performance of *maxRSSI* in room-level localisation for the two environments where we performed the experiment, and for the three beacon setups in the space, namely *centre*, *doors* and *far*.

A few interesting observations are in order: first, we verify that the device configuration within the space indeed matters. The performance of *maxRSSI* is largely variable between the various configurations for most users. For some users, performance might have less variability between configurations; our assumption is that this is related to which areas of the room each user covered when performing the experiment (recall that there was no prescribed path the users should follow). Second, it seems there might not be a global optimal configuration placement for all houses; we are only considering two environments in this paper and thus we cannot be conclusive, but the *centre* device configuration is dominant in the *demo-home*, in contrast with the *far* configuration which is the best performing in the *real-home*. Overall, the *maxRSSI* average accuracy is 67.77%.



(a) *demo-home* (b) *real-home*

Fig. 4. *maxRSSI* accuracies for all users and beacon configurations The accuracy of *maxRSSI* is calculated for three different device configurations (*centre*, *doors*, *far*) for each user. We observe significant variability in the *maxRSSI*'s accuracy depending on the configuration. Configurations behave similarly for all users in the same environment, i.e., configuration *centre* is best performing in the *demo-home* scenario (left), while *far* has the highest accuracy in the *real-home* scenario (right). We also observe a large variance across the users within a configuration set-up, even within the same environment, in particular for the *demo-home* case. Note that the axes start at 0.4 instead of 0, as *maxRSSI* exceeded this performance in all cases.

It is interesting to investigate whether we can attribute the low performance of *maxRSSI* to certain conditions. Figure 6a demonstrates the room-level locations estimated with *maxRSSI* for a case in the dataset where the *maxRSSI* approach is significantly low-performing: recID 19 – *centre* with just 57.03% accuracy. As is apparent, the recording includes long segments where no BLE RSSIs have been logged. It is also a case where two devices that are placed in separate rooms are sensed simultaneously, with the strongest RSSI value often heard from the adjacent room, and not the one that the user is in. We revisit this recording in Sect. 4, to demonstrate the improvements achieved in such a challenging case.

To conclude, with the **Room Detection** step, we have managed to exploit the IoT signals received in the smartwatch as the user is freely moving around the house, to infer estimates of their initial semantic locations. However, as we have seen, *maxRSSI*, can often have limited localisation accuracy, as it exploits the BLE data only and it thus unable to handle missing data or interferences from nearby rooms effectively. In the next section, we will see how we can make use of the additional mobility information provided by the smartwatch's IMU data, to eventually improve upon the results of *maxRSSI* in Sect. 4.

5 Motion Detection

The **Room Detection** step has provided initial room-level location estimates while the user is moving around the house, but has not taken into account any information about the mobility levels of the user in these rooms. The **Motion Detection** module aims to identify when the user is highly mobile, to help detect walking events that may hint room-transition, or idle events, which are correlated with the user not changing rooms.

5.1 IMU Data Preprocessing

A basic preprocessing was used to denoise, normalise and prepare the data for the **Motion Detection** module.

Similarly to the BLE data, we first perform linear interpolation to fill any missing values. Then, we perform moving median filtering, to achieve an initial reduction of noise, which is particularly present due to the sensor being on the user's wrist. A high-pass butterworth filter with cut-off frequency 0.3 Hz is then employed to remove the gravity vector from the acceleration data, and a low-pass butterworth filter with cut-off frequency 20 Hz (that corresponds to the range of frequencies of human activities) is used to further denoise the data [35]. The data are then bounded to $[-1, 1]$ and normalised to zero mean and unit standard deviation. The processed data can now be used from the **Motion Detection** module, as described below.

5.2 Algorithmic Details

For the **Motion Detection** module, we employ a simple *energyPeak* approach, based on the assumption that the overall energy of the acceleration signal recorded in the smartwatch should be higher during active phases, compared to idle phases.

There are many fitness-related smartwatch apps that aim to provide various information about mobility, including step counters or activity recognition apps, however these can be unreliable for the home environment case-study [6]. Smartwatches are known to effectively identify vigorous activities, such as switching between walking, cycling and running, but not so much activities performed at lower speeds [25] or combined with other domestic activities [6]. At the same time, smartwatch step counters commonly log many false-positive steps [23]; this might not be problematic for distances typically travelled outdoors, but it can be limiting when trying to estimate mobility within a house.

Regardless, for our scenario where floorplans of the house are also absent, a step-counter would have little benefit: as we are unaware of the size of the various rooms, as well as of their proximity, using a step counter as the mobility estimate would be inappropriate, as it would require assumptions about the number of steps that connect different rooms. It would also make it more difficult for the model to generalise across different houses.

Thus, we suggest that in order to have a system that easily applies to every environment and any type of smartwatch, and thus requires no calibration from the user, a much coarser mobility estimate would better suit our needs.

For our *energyPeak* approach, we first further filter the acceleration signal to the human walking frequencies, i.e., $[1.2, 2]$ Hz, to remove any redundant frequencies related to other activities. We then calculate the energy of the filtered signal through the Short-Time Fourier Transform (STFT) of the acceleration's magnitude, calculated in windows of $512 = 2^9$ samples, i.e., 5.12 s[4]. As we

[4] This ensures a synchronised 1-to-1 mapping between location and mobility estimates and explains the choice of 5.12 s window for the BLE RSSI data.

observed that high-energy contents gradually occur around ground truth transitions, we chose the envelope of the calculated energy as a final smoothed out energy estimate.

The motion estimates at time t are then chosen as the local peaks[5] of the normalised energy that are larger than the energy's average:

$$m_t = \begin{cases} \text{active}, & \text{if } acc_{\text{energy,smooth}_t} \text{ is a peak and} \\ & \quad acc_{\text{energy,smooth}_t} > \text{threshold}, \\ \text{idle}, & \text{otherwise.} \end{cases}$$

5.3 Evaluation

Our contributing dataset does not log any information about the activities performed during the experiment. As such, we cannot verify whether the *energyPeaks* we identified as walking events have been correctly classified as such. We can however verify how well our estimates capture ground truth room transitions. In a total of 555 transitions that occured across all the recordings in the dataset, our approach correctly identified 507 missing only 48, bringing the accuracy of *energyPeaks* for estimating room-transitions to 91.35%. A further 1388 timestamps were classified as walking events in a total of 45426 windows of not changing rooms; this is just 3.05% of the times the user stays within a room, which seems a valid behaviour of a user walking around a room while occupying it. In lack of ground truth activity data though, we cannot verify this any further.

To conclude, the suggested *energyPeaks* approach is simple and largely parameter free, yet effective at indicating room-transition events, albeit with some false-positives, as it cannot distinguish between movement during room changes, and casual walking within a room.

In the next section, we combine the outcomes of **Room Detection** and **Motion Detection** to provide a final, refined room-level user location.

6 Room Estimation

The **Room Detection** and **Motion Detection** modules have exploited the ambient IoT signals sensed from the smartwatch and its IMU to provide initial location estimates and estimates of walking events. In this step, we fuse this information together to improve upon the **Room Detection** module.

To fuse these two types of observations (r, m) - i.e., ambient location and mobility estimates - and thus derive refined room-level estimates, we take advantage of the inference capability of an appropriately defined Hidden Markov Model (HMM) with parameters $\theta \equiv (E_R, E_M, T, \pi)$, corresponding to the room emission probabilities E_R, the motion emission probabilities E_M, the state transition probabilities T, and the initial state vector probabilities π. We then refine our initial estimate by the parameters of the model that match our data for each house.

[5] Calculated with the `MATLAB`'s `findpeaks` function.

6.1 Model Definition

We define the state space of our HMM using "one hot encoding" for the set of all possible N^2 room transitions. We choose this room-transitions-as-a-state representation, as our mobility estimate m is concerned with likely room transitions. As such, a state at timestamp t is defined as:

$$s_{k,t} := (r_{i,t-1} \rightarrow r_{j,t}), \text{where } i, j \in \{1, 2, \dots N\}, \\ k := j + (i-1)N.$$

We also define the observation space using the sensor measurements obtained from the above modules, i.e., as the tuple $o_t := (r_t, m_t)$ at a given timestep. We assume that r and m are independent from each other.

To refer to a room-transition pair, and a combination of observations pair irrespective of time, we will use the following notation:

$$s_k \models (r_i \rightarrow r_j), \text{where } i, j \in \{1, 2, \dots, N\}, \\ k := j + (i-1)N,$$

and

$$o \models (r, m), \text{where } r \in \{r_1, r_2, \dots, r_N\}, \\ m \in \{idle, active\},$$

assuming that a state is always defined across two consecutive room locations in time, and a pair of observations is only properly defined when these occur simultaneously.

The emission probabilities for each type of observation are detailed in the following paragraphs.

Room Emissions. We initialise the room emissions E_R based on the assumption that when we are moving to a new room, we would expect the strongest RSSI to be emitted from the IoT device that is located in the room we arrive in. This is inclusive of the case when we stay in the same room in two consecutive time windows, as this event is encoded as a transition from a room to itself in the states of the HMM. Note that we have N rooms, but we can also observe the 'unknown room'; we can thus have $N + 1$ room observations.

The probability $p(r|s_k)$, i.e., the probability of observing room r when moving from r_i to r_j, is thus the following:

$$E_{r,k} = \begin{cases} 1 - \varepsilon_r, & \text{if } r_j \equiv r, \\ \varepsilon_r \ / \ N, & \text{otherwise.} \end{cases}$$

The above values are chosen to reflect our initial assumption that when moving into a new room, the beacon existing in this room will *almost surely* emit the strongest RSSI. The value ε_r is chosen to represent the uncertainty of the strongest RSSI being emitted from within the room of occupancy. For our experiment, we initialise $\varepsilon_r = 0.01$, to reflect the expected RSSI behaviour and for numerical stability.

Motion Emissions. The motion emission matrix must reflect that people are likely to be moving within a room without changing their semantic location, but not the opposite: it is impossible for a user to change their location without moving. As such, we define the probability of observing motion m when moving from r_i to r_j, i.e., $p(m|s_k)$, as:

$$E_{m,k} = \begin{cases} \begin{cases} \varepsilon_{m,c}, & \text{if } m : idle \\ 1 - \varepsilon_{m,c}, & \text{if } m : active. \end{cases}, & \text{if } s_k = (r_i \rightarrow r_j),\ i \neq j, \\ \begin{cases} 1 - \varepsilon_{m,s}, & \text{if } m : idle \\ \varepsilon_{m,s}, & \text{if } m : active. \end{cases}, & \text{otherwise.} \end{cases}$$

The values $\varepsilon_{m,c}, \varepsilon_{m,s}$ are chosen to model our prior on the user's mobility. Depending on the application, and on our knowledge of the environment, these can be chosen to be either identical in all rooms, or to reflect the different mobility patterns in each room (e.g., we might expect higher mobility in a kitchen compared to a study). In this work, we assign $\varepsilon_{m,c} = 0.01$ to reflect the certainty of movement during room-changes. In reality, this probability is equal to 0; there is no way to move from one room to another in the absence of a walking event. However, we chose to allow a very low probability of $E_{idle,i\neq j} = 0.01$ to avoid numerical instabilities in the code. To initialise the states representing a "stay-in-the-same-room" transition, we assign $\varepsilon_{m,s} = 0.3$, as it is still possible to walk within a room without any intention of changing location. Research has shown that people are expected to be sitted at least 75% of their time when working [14], so we define the above probability loosely based on these statistics.

Initial State Vector. The initial state probabilities vector π is defined by assuming we start uniformly at random from any state representing a stay-in-same-room transition:

$$\pi_k = \begin{cases} 1 \ / \ \mathrm{N}, & \text{if } s_k = (r_i \rightarrow r_i),\ i \in \{1, 2, \ldots, N\}, \\ 0, & \text{otherwise.} \end{cases}$$

Transition Matrix. To define the state transition matrix T of the HMM, we first need to remember that each state represents a *room transition* from room r_i to room r_j, i.e., $s_k \models (r_i \rightarrow r_j)$. The transition matrix T of the HMM is thus defined subject to the constraint that the first part of the next state must agree with the last part of the current state. Any state-transition that adheres to this constraint has the same probability. Any other state-transition has zero probability:

$$T_{fg} = \begin{cases} 1 \ / \ \mathrm{N}, & \text{if } s_f = (r_i \rightarrow r_j),\ s_g = (r_q \rightarrow r_p) \text{ and } j \equiv q, \\ 0, & \text{otherwise.} \end{cases},$$

where $f, g \in \{1, 2, \ldots, N^2\}$ and $i, j, q, p \in \{1, 2, \ldots, N\}$.

6.2 Inference

Having initialised the parameters $\theta \equiv (E_R, E_M, T, \pi)$ of the model, we can infer the most likely sequence of underlying states of the model for our observations, using the Viterbi algorithm. The trail of the landing rooms of these states corresponds to our room-level location estimates.

In principle, if the initial parameters of the HMM are carefully chosen, the Viterbi path can give us already improved location estimates, compared to the single sensor modality estimates. This is the simple version of our method, which omits the learning step of the HMM; we denote this special case of **RoomE** as Room Inference (**RoomI**). Figure 5 (second group of boxplots) summarises the performance of **RoomI** for the three different device configurations; in comparison with the **Room Detection** results from *maxRSSI* (leftmost group of boxplots), **RoomI** demonstrates an improved performance, averaging to 80.46%.

Figure 6b revisits the recording we studied in Fig. 6a. **RoomI** effectively fuses the location data with the mobility data to account for the no-signal phases of *maxRSSI*, and significantly improves on the adjacent-rooms' noise. We will see how we can further improve this behaviour by learning the parameters of the HMM, to allow for a more informed final decision on semantic localisation.

6.3 Parameter Learning

RoomI, which essentially traverses the initialised graph of our model, can already provide significantly improved semantic location information. In this step we aim to improve the model's parameters, by learning the optimal values that best describe the observed data. Once we have the learnt model parameters, we can use the updated model to infer a refined estimate of the user's room-level position.

One key detail in our approach is that we do not need to learn the full set of the model's parameters. This can be justified by recalling that we want to use the mobility data m to *correct* the **Room Detection** estimates; we thus need the model to expect a motion event when in states that represent room-transitions but not room-stays. As such, we can keep our assumption regarding whether an idle event or an active event should ignite a room transition fixed over time. We also need our model to be able to work its way through 'unknown room' occurrences, by only trusting the *energyPeaks* estimates in such a case; thus, the 'unknown room' observation should maintain its low probability occurrence at all times.

We thus train the room emissions $E_{r,new},\ r \neq$ 'unknown' and the state transitions T_{new}[6] only, with the Baum-Welch algorithm for expectation-maximisation (EM). The standard Baum-Welch algorithm is adapted as follows[7]:

For the E-step:

[6] Updating the state transitions might also be optional, as the transition matrix has been designed to only indicate allowed or not allowed transitions. However, learning this transition matrix can provide information about room connectivity.

[7] We refer the reader to the full set of Baum-Welch equations in [3].

– In the forward pass, the observed emissions can be modelled as the product of the corresponding elements of the two emission matrices, E_R, E_M, as these observations are independent:

$$\alpha_{t,k} = \sum_{f=1}^{N^2} \alpha_{t-1,f} T_{fk} E_{r_t,k} E_{m_t,k},\ k \in \{1, 2, \ldots, N^2\},$$
$$r \in \{r_1, r_2, \ldots, r_N, \text{“unknown room”}\},\ m \in \{idle,\ active\}.$$

where each forward probability $\alpha_{t,k}$ represents the joint probability of being in state $s_{k,t}$ at time t and observing the first t observations by time t, knowing the model parameters θ. It holds that $\alpha_{0,k} = \pi_k E_{r,k}(r_0) E_{m,k}(m_0),\ k \in \{1, 2, \ldots, N^2\}$.
– Similarly, the backward pass is defined as follows:

$$\beta_{t,f} = \sum_{f=1}^{N^2} T_{fk} \beta_{t+1,k} E_{r_{t+1},k} E_{m_{t+1},k},\ k \in \{1, 2, \ldots N^2\},$$
$$r \in \{r_1, r_2, \ldots, r_N, \text{“unknown room”}\},\ m \in \{idle,\ active\}.$$

with $\beta_{L-1,k} = 1,\ k \in \{1, 2, \ldots N^2\},\ L$: the total number of timestamps.
– The probabilities $\gamma_{t,k}$ of being in state $s_{k,t}$ at time t and $\xi_{t,f,k}$ of being in state $s_{f,t}$ at time t and state $s_{k,t+1}$ at time $t+1$, are defined as standard, subject to the above modifications of $\alpha_{t,k}, \beta_{t,f}$.

For the M-step:

– As we have discussed, we only need to learn the room emission probabilities. As such, each element $E_{r,k_{new}}$ in the updated emission matrix is the normalised expected number of times in state s_k and observing r, i.e.:

$$E_{r,k_{new}} = \frac{\text{expected \# times in } s_k \text{ and observing } r}{\text{expected \# times in } s_k},\ r \in \{r_1, r_2, \ldots, r_N\},$$
$$k \in \{1, 2, \ldots, N^2\}.$$

Note that the value for $r_i \equiv$ “unknown room” remains unchanged. We ensure that the full E_R matrix is always properly normalised, to ensure numerical stability and feasibility of the algorithm. The motion emission probabilities E_M also remain unchanged.
– To update the state-transitions, each element $T_{fg_{new}}$ in the updated transition matrix is the normalised expected number of transitions from s_f to s_g, i.e.:

$$T_{fg_{new}} = \frac{\text{expected \# transitions from } s_f \text{ to } s_g}{\text{expected \# transitions from } s_f},\ f, g \in \{1, 2, \ldots, N^2\}.$$

Refined Room Estimation. Following the training of these parameters, we perform inference on the learnt model, thus obtaining the final room-level location estimates generated by our method. There are two modes of inference function our method can consider: *offline* and *real-time* inference. In the *offline*

mode, the model computes the Viterbi path based on historical data (e.g., the full recording of the day), and then reports the full trail of estimated semantic locations of the user. In the *real-time* mode, the model computes the Viterbi using the L most recent observations ($L = 12$ in our case), and reports the current user location as the final room of the current Viterbi path estimate. Figure 5 summarises the performance of the *maxRSSI*, **RoomI** and **RoomE** estimates. **RoomE** achieves an average performance of 81.53%, an improvement of 20.3% on the *maxRSSI* estimates. **RoomE-RT**, denoting the *real-time* mode of our method, achieves an overall 77.6% localisation accuracy. As is apparent in the rightmost group of boxplots in Fig. 5, the *real-time* performance can be almost as good as the *offline* performance, depending on the device configuration. Figure 6c demonstrates the improvement of **RoomE** on the recording we examined in Fig. 6a and Fig. 6b.

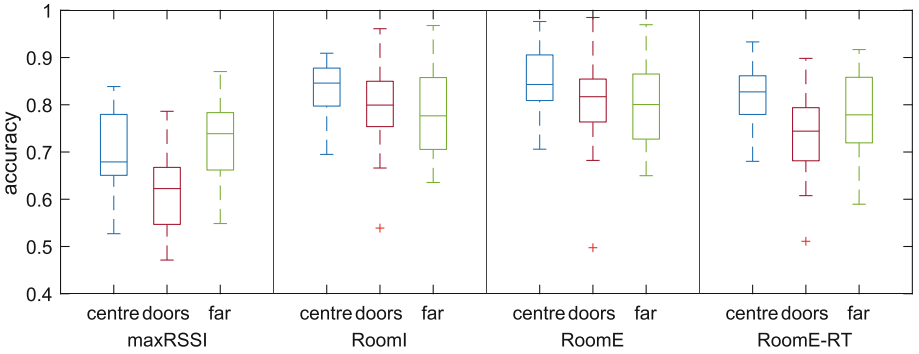
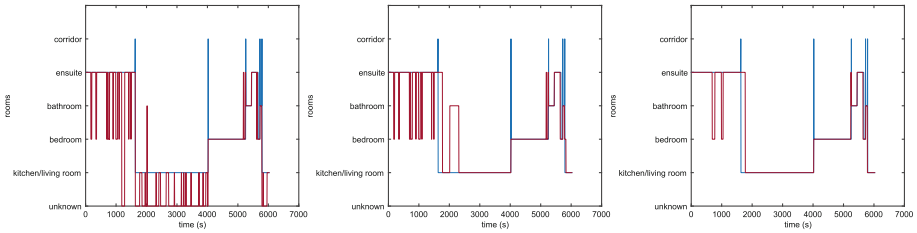


Fig. 5. Overall statistics The boxplots for the *centre*, *doors* and *far* configurations demonstrate the improved performance of **RoomI** compared to *maxRSSI* case, and the further improvement of **RoomE** against both **RoomI** and *maxRSSI*. The **RoomE-RT** section demonstrates the performance of **RoomE** in real-time. Note that the axes start at 0.4 instead of 0, as all methods exceeded this performance.

Semantic Map Estimation. With the learnt model, we can exploit the learnt state transition matrix to extract information about room connectivity, yielding a connectivity graph that resembles a semantic map of the house. To estimate the connectivity, we first estimate a room transition matrix from the state transition matrix, simply by averaging all the states that correspond to arrival in each room. We then calculate the semantic map as a normalised version of the room transition matrix, to remove the effect of how frequently each room is occupied: each line in the room-transition matrix is normalised, s.t. the probability of staying in the same room is 50%, and the probabilities of the remaining transitions sum up to the remaining 50%. In this way, we preserve the relative probabilities of transitioning to each room from a given starting point, but remove the information of how frequently each room is visited. This is the adjacency matrix of the possible room transitions, which can be compared to a ground truth adjacency matrix of the environment.



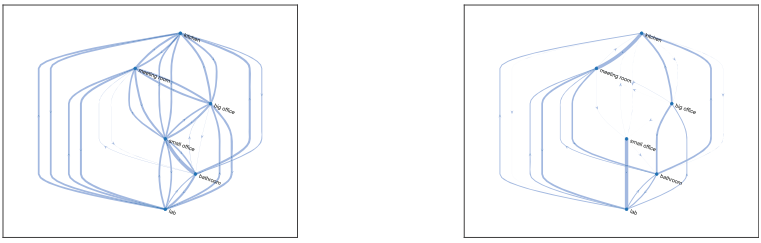
(a) *maxRSSI* (b) **RoomI** (c) **RoomE**

Fig. 6. Estimated vs ground truth locations for recID 19 : *centre*. (a) The *maxRSSI* fails to identify areas where the strongest signal comes from an adjacent room to the one the user is currently in. This is often the case for the adjacent `ensuite` and `bedroom` rooms of the *real-home*. For areas where there is no IoT device installed (e.g., corridors), or when no beacon is heard (e.g., in some parts of the *kitchen/living room*), the *maxRSSI* is unable to provide a location estimate. (b) **RoomI** effectively smooths out the previous unknown areas, and improves the noise from adjacent rooms, but is still prone to interferences. (c) **RoomE** further improves on the noise from adjacent rooms.

The average accuracy of the generated semantic maps is 97.85%. An original and estimated map of the *demo-home* environment are shown in Fig. 7

7 Related Work

Even though there has been extensive research on Wi-Fi positioning, Wi-Fi consumes significant power since the aim is to achieve signal propagation in long distances [34]. In indoor spaces however, this is not really required; with the



(a) Real map. (b) Estimated map.

Fig. 7. Semantic maps. The ground truth connectivity graph of the *demo-home* environment is compared to an estimated connectivity graph generated with our method. The estimated map resembles the original map in terms of connectivity and probabilities of transition between rooms. Some information regarding the frequency of visiting each room (i.e., the room occupancy statistics for the user) is still encoded in the estimated connectivity graph, leading to some missing edges, e.g., from `bathroom` to `big office`. These however correspond to low-probability transitions in the original graph, and so do not cause a significant error in our estimation.

development of the low-power energy efficient BLE, work has been done towards examining the potential of BLE beacons as Wi-Fi substitutes [9,34], showing that BLE both better relates RSSI to distance and improves localisation, compared to Wi-Fi. A number of approaches to Wi-Fi and BLE indoor positioning have since emerged; fingerprinting methods have traditionally been a classic approach [12, 16,26,30,37]. Though most methods try to tackle various problems arising from the fingerprinting technique, e.g., the size of the fingerprint map, or the map's dependency to the device, most fingerprinting methods require a map of the environment [1] and a well-defined path-loss model [2], which usually are both unknown in a home environment.

Significant work has also been developed using smart devices, such as smartphones, tablets and smartwatches. These are equipped with a number of inertial sensors, (e.g., accelerometer, gyroscope, magnetometer), and they can also work as Wi-Fi or bluetooth signal receivers. They are also often carried on a person for the majority of the day, meaning they can provide significant information about mobility [16,33]. IMU data from smart devices though are usually very noisy, as the device may be held in many different ways. In this context, estimating the direction of movement in particular can prove of significant difficulty. To address this, [22] propose a 3-step algorithm, which projects the walking direction in two directions: the direction of interference and the direction of the earth's magnetic field. Similarly, [15] and [32] try to address the problems of both step detection and heading estimation, by breaking down the step identification to the basics of human gait analysis, and exploiting projections to global and local coordinate systems. Robust-PDR (R-PDR) [32] in particular introduces an orientation correction algorithm (RIOT) to address drift in the heading estimation, by estimating the gravity vector through the acceleration data. RIOT exploits mainly the gyroscope, and its inherent property to have low drift and high reliability over a short time window. Hence, the relative difference between two consecutive time windows can provide accurate estimates of orientation change. Towards the same goal, [5] introduce deep neural networks (DNN) to minimise drift from IMU measurements.

The above methods try to solve the indoor localisation point through a geographical viewpoint. In the home scenario however, knowing the exact path or number of steps taken is of little interest to most applications. Instead, knowledge of the user's semantic position is informative enough, and helps to further preserve privacy by e.g., removing the need for personal details such as the user's height or leg length that are required for accurate PDR (step detection and step length), or for exact floorplans of the house as required by the map-matching approaches. In the area of semantic localisation of users, [21] introduced an algorithm based on human-robot coexistence to estimate the user's room-level location based on smartphone human activity recognition (HAR) and human-robot colocations. Interestingly, the authors did not record ground truth location of the user externally; instead, they installed BLE beacons in the rooms of interest, and assigned as ground truth location the location corresponding to the beacon that emitted the maximum RSSI value at any given time. A very

valuable approach from a human-only perspective, that is seamless and largely infrastructure-free is S-Smart [10]; this combines recognition of human activities with PDR to dynamically learn semantic positions of interest around the house, e.g., windows, doors, etc. It does require however a rather dense on-body sensor set-up, including a smartphone in the pocket, a wrist-worn IMU as well as a foot-mounted sensor, which are rather inconvenient for everyday use. On a similar note of using HAR to infer semantic locations around the house, the AtLAS system [19] fuses HAR and PDR from smartphone IMU data and Wi-Fi RSSIs to both locate landmarks in the environment and localise the user in it, but it requires Wi-Fi fingerprinting. In [28], the researchers, following the suggested approach of [4], are solely based on HAR estimated from smartwatch IMU data and BLE RSSIs from devices installed around the rooms of interest, but follow a supervised learning approach, to learn the correspondence between room-level locations and RSSI values.

RoomE can work around all the aforementioned usual problems of lack of environment maps, heading and distance drift due to noisy sensor data; the method learns the room-locations of the house over time in an unsupervised way, and is invariant to drift, as it does not seek to estimate exact trajectories. Instead, it utilises the energy levels of mobility instances and ambient signals to infer location and transitions, abstracting the notion of an exact trajectory to an accurate walk across nodes on a graph.

8 Conclusion

In this work we presented a comprehensive dataset, **WatchBLoc**, of ambient BLE and smartwatch IMU data for seamless room-level localisation. We discussed the challenges that occur in such positioning scenarios and demonstrated behaviours across different users, environments and device placements. We also introduced **RoomE**, a sensor fusion method acting on the above data to achieve privacy-preserving indoor localisation. Our method is less variable to device configuration compared to the state-of-the-art methods, and achieves accurate semantic localisation and mapping for the user at home.

There are still a few limitations in **RoomE**; we currently assume that every room in the house contains an IoT device that can be sensed through the smartwatch's IMU. In reality, there might be rooms in a house where no such device exists, as we do not yet truly live in fully-smart homes. Future work includes investigating how a minimal number of IoT devices can provide reliable localisation of the user both for rooms that are equipped with a device, but also for adjacent rooms that are not, extending our proposed approach to account for such scenarios.

References

1. Afyouni, I., Musleh, M., Basalamah, A., Tariq, Z.B.: Passive ble sensing for indoor pattern recognition and tracking. Procedia Comput. Sci. **191**, 223–229 (2021). https://doi.org/10.1016/j.procs.2021.07.028, www.sciencedirect.com/science/article/pii/S187705092101423X, the 18th International Conference on Mobile Systems and Pervasive Computing (MobiSPC), The 16th International Conference on Future Networks and Communications (FNC), The 11th International Conference on Sustainable Energy Information Technology
2. Bai, L., Ciravegna, F., Bond, R., Mulvenna, M.: A low cost indoor positioning system using bluetooth low energy. IEEE Access **8**, 136858–136871 (2020). https://doi.org/10.1109/ACCESS.2020.3012342
3. Bishop, C.M.: Pattern Recognition and Machine Learning (Information Science and Statistics). Springer, Berlin (2006)
4. Byrne, D., Kozlowski, M., Santos-Rodriguez, R., Piechocki, R., Craddock, I.: Residential wearable RSSI and accelerometer measurements with detailed location annotations. Sci. Data **5**(1), 180168 (2018). https://doi.org/10.1038/sdata.2018.168
5. Chen, C., Lu, X., Markham, A., Trigoni, N.: Ionet: Learning to cure the curse of drift in inertial odometry. CoRR abs/1802.02209 (2018), arxiv.org/abs/1802.02209
6. Chen, M.D., Kuo, C.C., Pellegrini, C.A., Hsu, M.J.: Accuracy of wristband activity monitors during ambulation and activities. Med. Sci. Sports Exerc. **48**(10), 1942–1949 (2016)
7. Christoe, M.J., Yuan, J., Michael, A., Kalantar-Zadeh, K.: Bluetooth signal attenuation analysis in human body tissue analogues. IEEE Access **9**, 85144–85150 (2021). https://doi.org/10.1109/ACCESS.2021.3087780
8. Faragher, R., Harle, R.: Location fingerprinting with bluetooth low energy beacons. IEEE J. Sel. Areas Commun. **33**(11), 2418–2428 (2015). https://doi.org/10.1109/JSAC.2015.2430281
9. Faragher, R., Harle, R.K.: An analysis of the accuracy of bluetooth low energy for indoor positioning applications (2014)
10. Hardegger, M., Roggen, D., Calatroni, A., Tröster, G.: S-smart: a unified bayesian framework for simultaneous semantic mapping, activity recognition, and tracking. ACM Trans. Intell. Syst. Technol. **7**(3) (2016). https://doi.org/10.1145/2824286
11. Hou, X., Bergmann, J.: Pedestrian dead reckoning with wearable sensors: a systematic review. IEEE Sens. J. **21**(1), 143–152 (2021). https://doi.org/10.1109/JSEN.2020.3014955
12. Jain, C., Sashank, G.V.S., N, V., Markkandan, S.: Low-cost BLE based indoor localization using RSSI fingerprinting and machine learning. In: 2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 363–367 (2021). https://doi.org/10.1109/WiSPNET51692.2021.9419388
13. Jiménez, A., Seco, F., Peltola, P., Espinilla, M.: Location of persons using binary sensors and BLE beacons for ambient assitive living. In: 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 206–212 (2018). https://doi.org/10.1109/IPIN.2018.8533714
14. Johansson, E., Mathiassen, S.E., Lund Rasmusse, C., Hallman, D.M.: Sitting, standing and moving during work and leisure among male and female office workers of different age: a compositional data analysis. BMC Public Health **20**(1), 826 (2020). https://doi.org/10.1186/s12889-020-08909-w

15. Kang, W., Han, Y.: Smartpdr: Smartphone-based pedestrian dead reckoning for indoor localization. IEEE Sens. J. **15**(5), 2906–2916 (2015). https://doi.org/10.1109/JSEN.2014.2382568
16. Lin, K., Chen, M., Deng, J., Hassan, M.M., Fortino, G.: Enhanced fingerprinting and trajectory prediction for IoT localization in smart buildings. IEEE Trans. Autom. Sci. Eng. **13**(3), 1294–1307 (2016). https://doi.org/10.1109/TASE.2016.2543242
17. McConville, R., Byrne, D., Craddock, I., Piechocki, R., Pope, J., Santos-Rodriguez, R.: A dataset for room level indoor localization using a smart home in a box. Data Brief **22**, 1044–1051 (2019). https://doi.org/10.1016/j.dib.2019.01.040, https://www.sciencedirect.com/science/article/pii/S2352340919300411
18. Mohammadi, M., Al-Fuqaha, A., Guizani, M., Oh, J.S.: Semi-supervised deep reinforcement learning in support of IoT and smart city services. IEEE Internet Things J. 1–12 (2017). https://doi.org/10.1109/JIOT.2017.2712560
19. Niu, X., Xie, L., Wang, J., Chen, H., Liu, D., Chen, R.: Atlas: an activity-based indoor localization and semantic labeling mechanism for residences. IEEE Internet Things J. **7**(10), 10606–10622 (2020). https://doi.org/10.1109/JIOT.2020.3004496
20. Roggen, D., et al.: Collecting complex activity datasets in highly rich networked sensor environments. In: 2010 Seventh International Conference on Networked Sensing Systems (INSS), pp. 233–240 (2010). https://doi.org/10.1109/INSS.2010.5573462
21. Rosa, S., Patanè, A., Lu, C.X., Trigoni, N.: Semantic place understanding for human-robot coexistence-toward intelligent workplaces. IEEE Trans. Hum.-Mach. Syst. **49**(2), 160–170 (2019). https://doi.org/10.1109/THMS.2018.2875079
22. Roy, N., Wang, H., Roy Choudhury, R.: I am a smartphone and i can tell my user's walking direction. In: Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services, pp. 329–342. MobiSys 2014, ACM, New York, NY, USA (2014). https://doi.org/10.1145/2594368.2594392, https://doi.org/10.1145/2594368.2594392
23. Saeb, S., Körding, K., Mohr, D.C.: Making activity recognition robust against deceptive behavior. PLOS ONE **10**(12), 1–12 (2015). https://doi.org/10.1371/journal.pone.0144795
24. Sakpere, W., Adeyeye-Oshin, M., Mlitwa, N.B.: A state-of-the-art survey of indoor positioning and navigation systems and technologies. South Afr. Comput. J. **29**, 145–197 (2017), www.scielo.org.za/scielo.php?script=sci_arttext&pid=S2313-78352017000300009&nrm=iso
25. Svarre, F.R., Jensen, M.M., Nielsen, J., Villumsen, M.: The validity of activity trackers is affected by walking speed: the criterion validity of garmin vivosmart(®) HR and StepWatch() 3 for measuring steps at various walking speeds under controlled conditions. PeerJ **8**, e9381 (2020)
26. Tegou, T., Kalamaras, I., Votis, K., Tzovaras, D.: A low-cost room-level indoor localization system with easy setup for medical applications. In: 2018 11th IFIP Wireless and Mobile Networking Conference (WMNC), pp. 1–7 (2018). https://doi.org/10.23919/WMNC.2018.8480912
27. Torres-Sospedra, J., Rambla, D., Montoliu, R., Belmonte, O., Huerta, J.: Ujiindoorloc-mag: a new database for magnetic field-based localization problems. In: 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN) pp. 1–10 (2015). https://doi.org/10.1109/IPIN.2015.7346763
28. Tsanousa, A., Xefteris, V.R., Meditskos, G., Vrochidis, S., Kompatsiaris, I.: Combining RSSI and accelerometer features for room-level localization. Sensors **21**(8) (2021). https://doi.org/10.3390/s21082723, www.mdpi.com/1424-8220/21/8/2723

29. Zsolt Tóth, J.T.: Miskolc IIS hybrid IPS: dataset for hybrid indoor positioning. In: 26st International Conference on Radioelektronika, pp. 408–412. IEEE (2016)
30. Wang, B., Zhou, S., Liu, W., Mo, Y.: Indoor localization based on curve fitting and location search using received signal strength. IEEE Trans. Ind. Electron. **62**(1), 572–582 (2015). https://doi.org/10.1109/TIE.2014.2327595
31. Weiss, G.M., Yoneda, K., Hayajneh, T.: Smartphone and smartwatch-based biometrics using activities of daily living. IEEE Access **7**, 133190–133202 (2019). https://doi.org/10.1109/ACCESS.2019.2940729
32. Xiao, Z., Wen, H., Markham, A., Trigoni, N.: Robust pedestrian dead reckoning (R-PDR) for arbitrary mobile device placement. In: 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 187–196, October 2014. https://doi.org/10.1109/IPIN.2014.7275483
33. Yang, Z., Wu, C., Zhou, Z., Zhang, X., Wang, X., Liu, Y.: Mobility increases localizability: a survey on wireless indoor localization using inertial sensors. ACM Comput. Surv. **47**(3), 54:1–54:34 (2015). https://doi.org/10.1145/2676430, http://doi.acm.org/10.1145/2676430
34. Zhao, X., Xiao, Z., Markham, A., Trigoni, N., Ren, Y.: Does BTLE measure up against Wifi? a comparison of indoor location performance. In: European Wireless 2014; 20th European Wireless Conference, pp. 1–6, May 2014
35. Zhao, Y., Yang, R., Chevalier, G., Gong, M.: Deep residual bidir-lstm for human activity recognition using wearable sensors. CoRR abs/1708.08989 (2017), arxiv.org/abs/1708.08989
36. Zhou, B., Li, Q., Mao, Q., Tu, W.: A robust crowdsourcing-based indoor localization system. Sensors (Basel) **17**(4), 864 (2017)
37. Zou, H., Huang, B., Lu, X., Jiang, H., Xie, L.: Standardizing location fingerprints across heterogeneous mobile devices for indoor localization. In: 2016 IEEE Wireless Communications and Networking Conference, pp. 1–6, April 2016. https://doi.org/10.1109/WCNC.2016.7564800