Published in final edited form as:
Proc ACM Int Conf Ubiquitous Comput. 2014 ; 2014: 159–170. doi:10.1145/2632048.2632097.

MobileRF: A Robust Device-Free Tracking System Based On a Hybrid Neural Network HMM Classifier

Anindya S. Paul[*,1], **Eric A. Wan**[*,1,2], **Fatema Adenwala**[2], **Erich Schafermeyer**[2], **Nick Preiser**[3], **Jeffrey Kaye**[4], and **Peter G. Jacobs**[*,1,3]

Anindya S. Paul: apaul@embedrf.com; Eric A. Wan: eric.wan@pdx.edu; Fatema Adenwala: adenwala@pdx.edu; Erich Schafermeyer: erich3@pdx.edu; Nick Preiser: preiser@ohsu.edu; Jeffrey Kaye: kaye@ohsu.edu; Peter G. Jacobs: jacobsp@ohsu.edu

[1]EmbedRF LLC, Portland, OR, 97201

[2]Electrical and Computer Engineering, Portland State University, Portland, OR 97207

[3]Biomedical Engineering, Oregon Health & Science University, Portland, OR 97239

[4]Neurology, Oregon Health & Science University, Portland, OR 97239

Abstract

We present a *device-free* indoor tracking system that uses received signal strength (RSS) from radio frequency (RF) transceivers to estimate the location of a person. While many RSS-based tracking systems use a body-worn device or tag, this approach requires no such tag. The approach is based on the key principle that RF signals between wall-mounted transceivers reflect and absorb differently depending on a person's movement within their home. A hierarchical neural network hidden Markov model (NN-HMM) classifier estimates both movement patterns and stand vs. walk conditions to perform tracking accurately. The algorithm and features used are specifically robust to changes in RSS mean shifts in the environment over time allowing for greater than 90% region level classification accuracy over an extended testing period. In addition to tracking, the system also estimates the number of people in different regions. It is currently being developed to support independent living and long-term monitoring of seniors.

Keywords

Indoor localization; indoor tracking; device-free passive localization; tag-free tracking; machine learning; neural network; health care; mobility

[*]Dr. Paul, Dr. Wan, and Dr. Jacobs have a financial interest in EmbedRF, a company that may have a commercial interest in the results of this research and technology. This potential conflict of interest has been reviewed and managed by OHSU.

INTRODUCTION

A number of different sensing technologies have been proposed to perform indoor localization ranging from RFID, infrared, ultrasonic transducers, RSS and ultra-wide-band (UWB) time-of-flight measurements. Typically, systems and approaches require the user to carry some type of physical device or tag. The focus of this paper is on the development of a *device-free* (or tag-free) system that works passively in the background and allows for indoor tracking without the person needing to carry any device. Applications where a device-free solution is desirable range from smart-home systems, security and intrusion detection, to virtual reality gaming. Our particular interest is in long-term health monitoring of seniors in support of independent living. Passive or ambient monitoring allows a caregiver or family member to observe deviations in patterns of activities of daily living while providing automatic notification if a change in health status or emergency event has occurred [24]. With traditional tag-based systems, seniors and especially those with cognitive decline, often forget or prefer not to wear their tags.

Previous approaches to device-free localization have included a variety of techniques and sensors. Video based tracking can often be effective, though performance may degrade with complicated background clutter and loss of privacy is a major concern for many applications. Simple contact switches or infrared (IR) motion sensors may be employed to determine room-level location. However, this approach does not provide accurate activity and mobility information and is inaccurate when more than one individual is in the living space. More accurate localization is possible using arrays of IR motion sensors, but such systems are expensive and complicated to install [6]. Ultra-wideband (UWB) or Doppler radar systems can be used to see through walls, although they require expensive equipment that limits their use in practice. A review of alternative sensors and approaches, including pressure sensors, load cells, thermal IR, ultrasound, and electric field capacitance, is provided in [13, 22].

The use of radio frequency (RF) attenuation is perhaps the most promising approach to device-free localization due to the availability of either existing Wi-Fi sensor networks or custom low-cost low-power transducers that provide measurements of RF attenuation in the form of received signal strength (RSS). Generally, the use of RF involves characterizing how a person affects the reflection or absorption of signals between multiple transmitters and receivers called *access-points (AP)*. The direct path between any two access-points is referred to as a *link* or *path*. Early demonstrations of device-free localization were performed in small highly controlled spaces [31, 39]. Subsequent publications in this area include [1, 9, 12, 15, 16, 22, 29, 34, 37, 40], as well as our own prior work [7, 33]. RF based methods can be categorized into *tomographic imaging*, *link-based*, and location based *fingerprinting*, as briefly reviewed below.

Radio tomographic imaging RF methods use hundreds of RSS measurement links typically by uniformly positioning access-points every meter or less around the perimeter of a room [1, 9, 11, 19, 38, 40]. An image map is created by measuring and modeling the attenuation between pairs of access-points. While these systems can be highly accurate in ideal

situations, they are usually based on line of sight (LOS) measurement models with accuracy degrading due to multipath in real-world living spaces.

Closely related to tomographic approaches are link-based RF localization methods that involve modeling the path-loss between pairs of access-point [10, 12, 22, 27, 35, 40]. Determining if a subject is on the LOS allows localization to be geometrically constrained to the link. Given sufficient number of links, variants of this approach have been quite impressive, allowing highly accurate localization and tracking using Bayesian approaches as well as methods for self-calibration of the link models [26, 27, 34]. Performance, however, is dependent on having a sufficiently large number of access-points and links (e.g., Ke, 2014, used 28 nodes with 378 links to cover a roughly 6.5m × 6.5m area, [12]). Concerns again exist over the impact of clutter in the home causing multipath and its effect on the path-loss modeling.

Another category of RF based approaches includes fingerprint-based methods. Similar to RSS device or tag-based localization, these methods involve a calibration phase in which a localization fingerprint or *radio-map* is learned by having a person stand (or walk) at a fixed number of known waypoints in the space. The advantage of this approach is that fewer access-points and signal paths are required, as a person does not need to stand directly on a link between two access-points. Multipath does not need to be analytically modeled and is implicitly accounted for in the data-driven calibration phase. Youssef et al., have developed a number of such systems over the years [15, 16, 18, 25, 26, 29, 39]. The Nuzzer system [29], for example is capable of better than 2 m location accuracy in a small office environment. The system is trained offline in which a person stands still for 60 seconds at 53 known locations spaced 2 meters apart while RSS data is recorded from only 6 paths. A Gaussian distribution is fit for the *mean* of the RSS data under an independence assumption for use in a discrete Bayes classifier. Spatial averaging is then used between regions to achieve better than 2-meter localization accuracy on standing data (results are not reported for walking data). The Nuzzer system also includes using variance thresholds to determine the number of people (up to 3) in different zones of the office. In Xu, 2013, the authors developed a system using *absolute* mean features with data collected for calibration using both standing and walking data [37]. The approach uses a conditional random field (CRF) to provide a Markov model of walking between regions. Experiments were performed in a $150m^2$ office with 13 transmitters and 9 receivers (119 links). Using a technique to sequentially cancel the effects of multiple people, tracking results were achieved for up to 4 people with an average location accuracy of 1.3m.

Our approach to device-free tracking is most closely related to the RF fingerprint methods. In our prior work [33] in a cluttered office environment using only 9 links with 5 access-points, we demonstrated region level localization (2–3 meter accuracy) and the ability to determine whether one or more people were in a room. Similar to the above-mentioned fingerprinting methods we used mean and variance features during standing and walking in order to train a neural network (NN), k-nearest neighbors (KNN), or Gaussian mixture model (GMM) classifiers. Subsequent to that work we were able to increase the resolution of our localization to less than 1m accuracy by using a finer grid of standing waypoints during training. However, a critical finding was that localization accuracy degraded

significantly over time, *i.e*., if a week or more separates the collection of testing and training data. Similar findings are reported in [14]. The issue is that *mean* RSS features that are relevant during *standing* phases vary considerably over time due to changes in the environment (e.g., a chair is moved), physical disturbances nearby (e.g., unknown changes in an adjacent office or apartment), or typical fluctuations in transceiver operation and un-modeled RF interference. Simply normalizing or subtracting by a background RSS mean when no one is in the room is not sufficiently effective at alleviating this issue due to the non-linear effects of RF multipath and signal transmission.

Our current approach addresses the robustness issue and adds additional system and algorithm capabilities. Instead of *localizing* the person we *track* the person during motion. The intuition is that if a person is standing there is high probability that they are not moving between regions and there is no need to localize them beyond their previously estimated location. We need only to track the sequence of changes between regions during motion. We use a hierarchical NN classifier to first determine whether there are 0, 1, 2, or more than 2 people in a room to be tracked. If the algorithm determines that only 1 person is in the room, we use a separately trained NN to estimate the probability whether a subject is standing or moving. Since localization is not performed during standing, short-term *mean* features that are prone to degradation are explicitly removed from use and only short-term variance features used (other features are also compared for robustness). Further we use a hybrid NN-HMM classifier to track the sequence of motion between different regions and between standing versus walking. The HMM model of motion assumes equal probability of a person moving to adjacent regions or staying in the same region. The fixed probability transitions are then adjusted by blending with the classification probability of a person standing or moving. A second NN classifier provides the posterior probability of being in a specific region given RSS features. Finally, a Viterbi algorithm is used to determine the maximum-likelihood sequence of regions that a person moves through in the environment.

As described above, the system is capable of classifying whether there are 0, 1, 2, or more people in larger regions (room level). This is useful for monitoring caregiver and social interaction, as well as for automatically parsing single-person data for long-term studies. The NN-HMM approach is then used only when 1 person is in the room for more accurate localization and tracking. In addition, we use the same passive RSS access-points to accurately determine walking speed. Walking speed and changes in walking speed have been shown to be important health indicators and may be used in detecting the onset of cognitive decline. For estimating walking speed we use both the timing and shape of the RSS as detected between two links placed across a hallway or using a single link in a doorframe. Walking speed estimation is better than 8 cm/s root-mean-square error (RMSE) relative to a GAITRite gait mat used for ground truth. Details on the walking speed component can be found in our recent publication [8].

Our new approach to device-free localization is also being developed as part of a complete indoor mobility assessment and tracking system, called MobileRF, in collaboration with the company EmbedRF. MobileRF includes a tag-based multi-patient mobility assessment mode in which the tag is worn by the patient(s) and localization performed using ultra-wideband (UWB) ranging in combination with sigma-point Kalman filtering (SPKF) for tracking and

simultaneous localization and mapping (SLAM) for self-calibration [7, 23]. Additional inertial sensors (3D gyroscopes and accelerometers) allow for more detailed assessment of gait features. The tag-based mode can also be used for ground-truth during the calibration phase of the device-free approach. In this paper, we focus on describing only the device-free solution for performing indoor tracking.

METHODS

In this section, we describe the hardware and software data acquisition system, the facilities and procedures used for testing, and the tracking algorithm design.

Hardware

Wall-mounted access-points (APs) correspond to miniature transceiver modules manufactured by EmbedRF (see Figure 1). The APs are configured to wirelessly communicate with each other as well as to a central hub (also containing an EmbedRF transceiver). The hardware consists of a small 8-bit microcontroller (PIC 16F690), a Texas Instruments radio transceiver (CC1101), and some passive components. RF energy inside the room absorbs and reflects differently depending on where a person is located. Any motion of a person through the room disrupts the radio field, which changes the RSS readings recorded by the access points and hub.

Testing facilities and experiments

The experiments and results reported here were performed in two different locations. The first site was the OHSU Point-of-Care Laboratory (PoCL) (see Figure 2 floor plan). The PoCL is a simulated apartment consisting of three rooms, a bedroom, bathroom and combined kitchen / living room; the space is filled with furniture and appliances typical for a home environment including a kitchen sink, dish washer, couches, television, bed, and bathroom with toilet and shower. The small apartment is approximately 9m × 7.5m in size and had 9 APs and 1 hub installed resulting in 17 RSS links. The PoCL was divided into M=7 regions for tracking (see Figure 2). The second site used to further confirm performance was the Portland State University Biomedical Signal Processing Laboratory (BSPL). The BSPL is a multi-purpose electronics laboratory and open office space of dimension 5m × 4m. In the BSPL, 5 APs and 1 hub were used resulting in 9 RSS links; M=4 quadrants were used for testing tracking (floor plan is not shown due to space limitations).

The RSS links for the PoCL are also illustrated as arrows in Figure 2. Note that not all AP to AP RSS paths were measured because the system is designed to have all access points run at ultra-low power levels on a single set of AA batteries for over a year. To achieve this long battery life, the transmission of the signal is started by the *initiator*, which sends a signal to all other APs and directly to the hub. The APs receive this signal and then transmit directly to the hub. The initiator transmits data every 250 milliseconds (ms). Each of the APs including the initiator is asleep for most of the time except when they wake up periodically to listen and then transmit. The transmission is done at 915 MHz to avoid interference from Wi-Fi or other 2.4 GHz and 5 GHz signals typical in a home environment and the data bytes transmitted are kept to a minimum of 10 bytes per payload, again to minimize battery drain.

The power drain of the system is an order of magnitude less than ZigBee and Wi-Fi based systems.

An example of the raw RSS outputs as a person moves through the PoCL is shown in Figure 3. It is important to note that the disruption in the RSS does not require a person to be directly in the line-of sight (LOS) between two access-points. Multipath and other disturbances are actually beneficial for this approach.

Software data acquisition system

Custom software was used for the data collections running on a PC computer and written using C#.NET. The software collected the transceiver data using a hub connected to the laptop USB port. A user can use the software to designate *regions* that are used during tracking, *waypoints* that are specific locations within a region, as well as general *activities*, e.g. sitting, standing, or walking. Different *scripts* can then be run, which automatically direct the participant using a text-to-speech module through a specified sequence of instructions. A typical calibration script instructs the person to move to region 1 (*entry area*) and walk around, then move to way point 1 and stand there, then move to region 2 (*TV area*) and move around etc. until the user has moved to all of the required regions and waypoints. The scripts are fully customizable and can direct a person to perform other activities of daily living, e.g., opening and closing the refrigerator, washing dishes, sitting on the couch, or walking paths at different speeds through specific sequences of waypoints. The RSS data acquired during this calibration script are time aligned with the instruction labels and then can be used to train or test the classifier and tracking system as described further below.

Pre-processing and feature extraction

The raw RSS information acquired at a 250 ms sample rate for each link serves as the data input to the tracking system. An example plot of the RSS for each link is shown in Figure 3.

Critical to classification performance is finding feature representations for the data that allows for high discrimination while maintaining high generalization between users and over time. Raw RSS data signals are first converted from dBm into power (mWatts):

$$RSS_w = 10^{RSS_{dbm}/10} \quad (1)$$

Next, “background subtraction” is performed in which the long-term mean of the background RSS (when no one is in the space) is subtracted from the raw RSS power values. The background-subtracted data contains information about RSS *variability* due to the person’s presence inside the room, while attempting to eliminate other variability due to room geometry, furniture location, and static multipath. Note that this subtraction only *partially* eliminates the effects of static changes in the space due to the inherent nonlinear nature of RF path-loss and multipath effects. This is a key point. While it is intuitive that a person’s specific location in a room will cause a mean shift in the observed RSS level, the absolute *amount* of this shift may not be consistent over time due to many factors. We found that on numerous occasions, the mean RSS level for one or more links would suddenly shift and/or drift in level during the course of a single trial. It is unlikely that this was a hardware

or data acquisition issue with the transceiver (Texas Instruments CC1101) as there are no reports on this being typical of the part. These shifts are likely caused by environmental factors or outside RF interference. Mean shift are even more problematic during longer term studies involving over a month of time, where changes in furniture location (simply moving a dining room chair or leaving a door open in a different position) can cause changes in both the absolute background RSS level as well as the *relative* shift due to a person's presence in a specific location. Figure 4 demonstrates instances of mean level shifts of RSS for a) when a chair is moved, and b) over a 10 hour period.

In our preliminary work, we investigated use of short-term mean RSS features. Several experiments were performed in which subjects were instructed to stand at fixed waypoints spaced on a grid of less than 1m separation. Collections were performed in a cluttered office, empty room, and outdoors. Our findings were that when using only mean features, sub-meter accuracy in localization (with greater than 80% region level accuracy) was only possible in the ideal case when training and testing data was collected immediately following each other. Within a month of separation, accuracy fell to well below 50%.

Because of these factors, we concluded that using the *mean* RSS level as a feature is problematic for our long-term study. Our application goal was to develop a robust system that can be deployed in a person's home with minimal calibration and re-calibration. In the current system, we addressed this issue in two ways, 1) we used only RSS variance features and we explicitly did not use mean level RSS features in our classification procedure, and 2) we developed a hybrid NN-HMM classifier to track the sequence of motion and avoided explicit localization when a person was determined to be standing still. Specifically, our primary feature used for tracking was the short-term (5 second (s) window) variance of the RSS signal. A *sliding* window was used such that the variance features were updated at every 250 ms. The full feature vector then consisted of the concatenation of all variance features for each RSS link. This relatively simple change in the feature set had a significant effect on our overall results as will be described later. Note that we also investigated different time windows for calculating the variance (1 s, 2 s, etc.) as well as short-term Fourier Transform based spectral features, Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA); however, we did not find any significant advantage over using the simple 5 s sliding window variance.

Classification and tracking algorithms

In this section, we describe a hierarchical tracking algorithm that works by (1) first estimating the number of people in a room, (2) next identifying whether a person is standing or walking, and (3) lastly tracking the person using a hybrid NN-HMM. Steps 1, 2, and 3 all utilize different NN classifiers that must be trained on data acquired during a calibration run. Both the NN and HMM designs are described below.

Neural network design and training

Three separate feed-forward NN classifiers were trained and used in the system, (1) NN_p estimates the probability of 0, 1, 2, or more than 2 people being present in a room, (2) NN_s estimates the probability of whether a person is standing still or walking, and (3) NN_r

estimates the probability of a person being in a region in the room. We define the following notation. Define R_i to be the i^{th} region, $R_1 \leq R_i \leq R_M$, where M is the total number of regions in the space. Let $\mathbf{x}_t = [RSS_w(1), RSS_w(2), \ldots. RSS_w(N)]$, be the vector of all RSS measurements at time t after mean background subtraction, where N is the total number of links. The vector of short-term variance features (5 s sliding window) derived from $\mathbf{x}_t$ at time t is denoted as $\boldsymbol{\theta}_t$. Calibration data, which involves the known number of people, standing versus walking, or region location, was used to train each of the 3 NNs. During training, the hidden layer structures of the classifiers as well as the appropriate weights were determined using 10-fold cross validation. Additional specifics on the calibration data used and network structure for each NN is provided in the Results section.

Once trained, a NN takes a feature vector, $\boldsymbol{\theta}_t$ as input and produces an output that can be interpreted and used as posterior probabilities, specifically, the first network,

$$\mathrm{NN}_p(\boldsymbol{\theta}_t) \approx p(p_{j,t}|\boldsymbol{\theta}_t) P_0 \leq p \leq P_{>2} \quad (2)$$

gives the probability of detecting the number of people, $p=P_i$, given $\boldsymbol{\theta}_t$ at time t. The next network,

$$\mathrm{NN}_s(\boldsymbol{\theta}_t) \approx p(s_{k,t}|\boldsymbol{\theta}_t) s=S_0, S_1 \quad (3)$$

gives the probability of standing ($s = S_1$) vs. walking ($s = S_0$) at time t. Finally,

$$\mathrm{NN}_r(\boldsymbol{\theta}_t) \approx p(r_{i,t}|\boldsymbol{\theta}_t) R_1 \leq r_i \leq R_M \quad (4)$$

gives the probability of being in region $r = R_i$ at time t given $\boldsymbol{\theta}_t$. A hard classification can be made by selecting the maximum output for a given classifier. Specifically,

$$\hat{R}_t = \arg \max \left[p(r_{i,t}|\boldsymbol{\theta}_t) \right] \quad \hat{P}_t = \arg \max \left[(p_{j,t}|\boldsymbol{\theta}_t) \right] \quad \hat{S}_t = \arg \max \left[(s_{k,t}|\boldsymbol{\theta}_t) \right]. \quad (5)$$

Note that in practice we only make a hard decision for the first classifier that determines the number of people. The hybrid NN-HMM uses the probability estimates from NN_s and NN_r to perform tracking as explained subsequently.

We compared alternative classifiers, including k-nearest neighbor, Gaussian mixture models, and support vector machines. All yielded comparable results, however, the standard NN provided slightly better performance and had the advantage of easily providing posterior probability estimates. The NN is also computationally and memory efficient to implement after training for real-time use.

Hybrid NN-HMM design

The HMM performs maximum likelihood sequence estimation by combining the probability of standing vs. walking (output from NN_s) and the probability of the person being in a region (output from NN_r) with a probabilistic transition model of a person moving from one region to another. A Markov-model is assumed, in which the hidden states are the regions.

The transition probability matrix $\boldsymbol{A}_{ij}$ is denoted as the probability of moving from region r_j to r_i

$$\mathbf{A}_{ij}=\{a_{ij}\}=p(r_{i,t}|r_{j,t-1})R_1 \le r_i \le R_M \text{and} R_1 \le r_j \le R_M. \quad (6)$$

The observations emitted from each state of the HMM model are the features θ_t. The observation probability is derived from $\text{NN}_r(\theta_t)$ and denoted as the likelihood of observing the feature vector θ_t at time t given a region r_i

$$\boldsymbol{B}(\boldsymbol{\theta}_t)=\{b_i(\boldsymbol{\theta}_t)\}=p(\boldsymbol{\theta}_t|r_{i,t})R_1 \le r_i \le R_M. \quad (7)$$

The initial probability of starting a walk from a region R_i is given as $\pi_i = p(r_i)\ R_1\ \le r_i < R_M$.

For the transition matrix, we assumed that the person was equally likely to move to any of the regions surrounding the current region as well as staying within that region. For example, if there are 4 directions a person can move from a given region, the transition probabilities to each region will be 1/5 since there is also a probability that they will stay within that region. The states and transitions probabilities for the PoCL are illustrated in Figure 5. Note that future work might involve learning the transition probabilities from observed patterns of a person walking over time.

The above transition matrix and probabilities assume a person is *moving*. Whereas a traditional HMM uses fixed transition probabilities, we dynamically adjust the probabilities based on the NNs output that determines the probability of a person standing versus walking. Thus we redefine the region transition probabilities at time t by combining the state transition matrix $\boldsymbol{A}_{ij}$ conditioned on whether a person is standing versus walking,

$$A'_{ij}=p_{(s=S_1),t}p(r_{i,t}\left|r_{j,t-1}, s_t=S_1\right.)+(1-p_{(s=S_0),t})p(r_{i,t}\left|r_{j,t-1}, s_t=S_0\right.)=\left[p_{(s=S_1),t}(\boldsymbol{\Delta}_{ij})+(1-p_{(s=S_1),t})\mathbf{A}_{ij}\right]; \quad (8)$$

where $p_{(s=S_1),t}$ is denoted as probability of standing at time t, i.e., the output of NN_s, and,

$$\begin{aligned}\boldsymbol{\Delta}_{ij} &= 1 \forall R_i = R_j \\ &= 0 \text{elsewhere}.\end{aligned} \quad (9)$$

implies that the probability equals 1 of staying in the same location given a person is standing. Similarly, observation probability $\boldsymbol{B}(\boldsymbol{\theta}_t)$ is modified to obtain $\boldsymbol{B}'(\boldsymbol{\theta}_t)$,

$$\boldsymbol{B}'(\boldsymbol{\theta}_t)=\left\{b'_i(\boldsymbol{\theta}_t)\right\}=p_{(s=S_1),t}(1/M)+(1-p_{(s=S_1),t})\boldsymbol{B}(\boldsymbol{\theta}_t) \quad (10)$$

which implies the region based classifier is less informative (equal probabilities) when a person has high likelihood of standing.

The output of the NN-HMM is the maximum-likelihood sequence of regions,

$$\left\{\hat{R}_0^{HMM}, \hat{R}_1^{HMM}, \ldots \hat{R}_T^{HMM}\right\} = \arg\max_{r_0 \ldots r_T} \left[p(r_{i,t}|\boldsymbol{\theta}_t)\right]. \quad (11)$$

Applying Bayes formula,

$$p(r_{i,t}|\boldsymbol{\theta}_t) \propto p(\boldsymbol{\theta}_t|r_{i,t})p(r_{i,t}). \quad (12)$$

where $p(r_{i,t})$ is expressed recursively,

$$p(r_{i,t}) = \sum_{j=1}^{M} p(r_{i,t}|r_{j,t-1})p(r_{j,t-1}) \quad (13)$$

using the transition probabilities defined earlier. A Viterbi search algorithm is then used to estimate the most likely path of the person moving through a set of regions. The pseudo-code to compute $\hat{R}_t^{HMM}$ using the NN-HMM method and Viterbi search is given below:

Initialization:

- $\pi_i = 1$, $R_1 \leq r_i \leq R_M$ assuming all regions are equally likely at the start of the walk.
- Viterbi initialization:

$$\delta_1(i) = \pi_i b_i'(\boldsymbol{\theta}_1) \quad 1 \leq i \leq M$$

$\psi_1(i) = 0$

where $\boldsymbol{B}'(\boldsymbol{\theta}_1)$ is obtained from $\mathrm{NN}_r(\boldsymbol{\theta}_1)$

$$\boldsymbol{B}'(\boldsymbol{\theta}_1) = \left\{b_i'(\boldsymbol{\theta}_1)\right\} = p_{(s=S_1),1}(1/M) + (1 - p_{(s=S_1),1})\boldsymbol{B}(\boldsymbol{\theta}_1)$$

- Recursion:

  while $t \leq T$

  - Run NN_p and obtain $p(p_{j,t}|\boldsymbol{\theta}_t)$ so that $\hat{P_t} = \mathrm{argmax}[p(p_{j,t}|\boldsymbol{\theta}_t)]$ $P_0 \leq p \leq P_{>2}$

    if $\hat{P_t} = 1$ (single person case)

    - Run NN_s and obtain $p(s_{k,t}|\boldsymbol{\theta}_t)$ $s = S_0, S_1$
    - Compute probability of standing $p_{(s=S_1),t} = p(s_t = S_1|\boldsymbol{\theta}_t)$
    - Compute new transition matrix

      $$\mathbf{A}_{ij}' = p_{(s=S_1),t}(\boldsymbol{\Delta}_{ij}) + (1 - p_{(s=S_1),t})\mathbf{A}_{ij}$$

    - Run NN_r to obtain $\boldsymbol{B}(\boldsymbol{\theta}_t)$

$$\boldsymbol{B}'(\boldsymbol{\theta}_t)=\left\{b_i'(\boldsymbol{\theta}_t)\right\}=p_{(s=S_1),t}(1/M)+(1-p_{(s=S_1),1})\boldsymbol{B}(\boldsymbol{\theta}_t)$$

- Viterbi recursion:

$$\delta_t(i)=\max_{1\le j\le N}\left[\delta_{t-1}(j)a_{ij}'\right]b_i'(\boldsymbol{\theta}_t)$$

$$\psi_t(i)=\arg\max_{1\le j\le N}\left[\delta_{t-1}(j)a_{ij}'\right]$$

$1 \le i \le M$

- $t = t + 1$

else (zero person and multi-person case)

- Reject the data
- $t=t+1$ end

end

- Termination:

$$P^*=\max_{1\le j\le N}\left[\delta_T(j)\right]\quad \hat{R}_T^{HMM}=\arg\max_{1\le j\le N}\left[\delta_T(j)\right]$$

- Viterbi back-tracking:
  - Optimal sequence of regions can be obtained by backtracking

$$\hat{R}_T^{HMM}=\psi_{t+1}(\hat{R}_{t+1}^{HMM})t=T-1,T-2,\ldots$$

RESULTS

A number of subjects performed several experiments over multiple days and weeks at each location to calibrate and test the accuracy of the different classifier components and overall hybrid NN-HMM tracking system. Results on multi-person classification, standing versus walking, and region based tracking are presented below.

Single vs. multi-person detection

The first level classifier NN_p discriminates between 0, 1, 2 or greater than 2 person occupancy. In the PoCL, data was collected for 30 minutes in which either 1, 2 or 3 people were asked to walk around the entire location. When 2 or 3 people were present, normal interactions were simulated whereby the subjects walked in different rooms, walked in the

same room, followed one another, stood still, etc. Classification performance was evaluated using independent datasets for training and testing. The number of hidden nodes for NN_p determined using cross validation was 5 for both PoCL and BSPL. Classification accuracy was about 90% using the RSS 5 s sliding window variance features. Table 1 shows classification performance in terms of a confusion matrix for the PoCL. Note that determination of 1 vs. multi-person was performed at a 4 Hz rate. However if a full 5 minute segment of data was used to make the estimation based on a simple *majority vote* of the 4 Hz classifier decisions, the accuracy increased to 100%.

A second test on multi-person detection was also performed in the BSPL, in this case to determine whether it was possible to discriminate both the number of people (1 or 2) and in which of 4 regions each was located. We found that our classifier was able to perform the task successfully 92% of the time using the 5 s sliding window.

Standing vs. walking classification

Once the NN_p has determined that only one person is in the space, the NN_s determines whether the person is standing or walking. Data were collected in the PoCL and BSPL in which a person either stood at a waypoint or walked in each region for approximately 30 s. Again, the 5 s sliding windowed variance features were used as input to the classifier, which made a decision at a 4 Hz rate. The number of hidden neurons was 3 as determined by performing cross validation on the PoCL and BSPL data. Table 2 demonstrates the performance based on a confusion matrix. Classification accuracy was over 95%. Only classification results for the PoCL are shown, as classification accuracy was approximately the same at the BSPL. Note that in practice the use of this classifier is not a “hard” decision, but rather to estimate the probability that a person is either standing or walking for use in hybrid NN-HMM tracking.

Region-level tracking

The full hybrid NN-HMM tracking system combines both the output of the standing vs. walking classifier NN_s and the region based classifier NN_r. The NN_r was trained by using calibration data collected in the PoCL or BSPL in which a person walked randomly in each region for approximately 60 s. The number of hidden nodes was determined to be 20 and 10 for the PoCL and BSPL NN_r, respectively. The HMM incorporates the transition model and is used to perform maximum-likelihood sequence estimation to obtain the most likely path the user has taken. The classifier performance is expressed based on percentage of time the estimated region matched the true region. Note that the dimension of each region was on average 2.7m × 2.7m. Data collection and feature representation was as before.

Total performance averaged over all regions is shown for both PoCL and BSPL in Table 3. To evaluate robustness, this table also shows performance when training and testing were performed on the same day, separated by a week, and separated by a month. Table 4 and Table 5 provide confusion matrices whose diagonal elements show accuracy of each region and off-diagonal terms exhibit misclassification. Table 3 also includes performance comparisons when mean features were used with and without the variance features. Note that same NN_r calibration dataset and 5 s sliding window as variance features were used to

compute mean features. As evident, classifier performance when using only short-term variance features did not significantly degrade over time. In contrast, the use of mean features resulted in significantly poorer performance. The test facilities at PoCL and BSPL were actively used by other research groups for experiments in between our evaluations. This resulted in the movement of furniture, bookcases, and overall floor-plan layout on numerous occasions. Also there were multiple instances when APs were taken down from walls and then replaced between data collections. Over the course of a month, the hub AP was moved several feet, which may account for some of this performance loss over the period. Overall, system tracking remained robust (when using only the variance features), and accuracy was sufficient for health monitoring.

In Figure 6 we show how the system tracked a person moving sequentially from region 1–7 at PoCL. The red trace is the ground truth while the blue trace is the location estimated by the NN-HMM tracking algorithm. Performance was good both on same day training and when separated by a month. Note that when errors were made, the tracking usually quickly recovered, indicating that post-processing involving smoothing may be possible to further improve performance. We also performed experiments to evaluate the advantage of using the hybrid NN-HMM over simply using the single neural network NN_r classifier to perform region level localization. Figure 7 illustrates the difference in tracking performance over the sequence of regions. Table 6 compares average performance over all regions and also includes results when using mean features for comparison. As is clear from these results, performance of the hybrid NN-HMM classifier improves significantly when incorporating prior knowledge of a person's state within a room as well as blending the probabilities of standing and walking. The advantage of using variance features instead of mean features is also again evident.

In another evaluation, we compared effects of training/testing on a) the same person (same as Table 3), b) using 75% less training data, and c) training on one person and testing on a different person. Results are summarized in Table 7 which demonstrate another form of robustness of our approach. This is important for our application of tracking and mobility monitoring of the elderly, where a different person other than the senior might perform the initial calibration during set-up. A short calibration routine is desired to save on installation time. We also investigated the effects of reducing the number of APs. In the PoCL, we reduced the number of APs from 9 to 5 resulting in 8 fewer RSS links (see Figure 2). In this case, region level classification accuracy only reduced by 7% for same day training and testing and only 4% when separated by a month. This shows the ability to trade off the number of APs used for performance accuracy.

The final set of evaluations involved tracking a person as they walked naturally in the space following a pre-defined path. Continuous walks are more challenging to track as subjects spend only a few seconds at each region mimicking typical movement of a person inside their home. An example walking path at PoCL is displayed in Figure 8. In the example shown, we also increased the number of regions from 7 to 13 (some of the regions are now sub-meter dimension, with the average grid size being 1.6m × 1.6m). While the overall region level classification accuracy decreased by 6% relative to using 7 regions on same day training and testing, the estimated walking path still closely estimates the true path (Figure

9). Note that while we are using a 5 s window to calculate the variance features, this is a *sliding* window, which in combination with the HMM transition constraints allow us to accurately track a person walking through a given region even if they spend less than 5 s crossing the region. Additional testing involved walking in the reverse direction, at different speeds, as well as stopping and standing occasionally along the path. Performance was similar in all cases further indicating the robustness of the method.

Additional system capabilities and results

As discussed earlier, our tracking system is designed for use as part of an overall passive system for monitoring the elderly. Using similar neural network classification techniques, experiments were performed to discriminate between standing versus sitting on a couch. On average, we found classifier accuracy to be above 98% for same day train and test and around 92% when train and test were one week apart. Furthermore, using timing information between RSS link crossings, we were able to estimate walking speed at an RMSE of less than 8cm/s relative to a gait mat used for ground truth. The details of the experiments and results are demonstrated in a separate publication [8].

DISCUSSION AND CONCLUSIONS

In this paper, we have presented a novel solution to passive device-free tracking based on detecting the RSS disruptions between wall-mounted access-points. Whereas most in-home tracking and monitoring systems are based on wearable technology, our system does not require the user to carry any device or tag. One novelty of the presented algorithm is that it takes advantage of the dynamic changes in RSS during movement to intelligently track a person only when they are moving. Several classifiers are implemented. A first level neural network classifier determines the number of people in a room or region. When a single person is detected, a hybrid NN-HMM is used to optimally track the sequence of movements of the person from region to region. The algorithm is robust to long-term drift in the RSS and other changes in the environment (e.g., moving furniture) that might affect signal multipath. The approach specifically excludes data when more than one person is present in an environment, thereby making it optimal for performing movement monitoring for seniors either living alone or with other people. Determination of the number of people, classification of standing versus walking, and region level classification were all over 90% accurate. Experimental results further demonstrate reliable region level tracking to within approximately 1–2 m. We report results using a moderate number of low-cost APs.

Seifeldin et al. [29] have also presented work using a small number of APs within a real world environment to localize people. Results presented by this group demonstrate localization accuracy (< 2 m) using standard Wi-Fi RSSI radio-maps and remarkably using only 3 APs. This group has also shown that they can localize multiple stationary people within a region both within this paper and they demonstrate multi-entity localization in Sabek et al. [25]. The primary difference between our work and these papers is that they had trained and tested on people only when they were stationary whereas we train and test on both stationary and moving people. So while the Seifeldin and Sabek papers describe an algorithm that accurately localizes stationary objects, they are not tracking user's sequence

of movements. Furthermore, this group does not demonstrate algorithm performance over long time periods during which RSS can drift. Moving objects and people affect the RSS differently than stationary objects; we have reported here that when tracking objects that alternate between moving and stationary, it is important to first estimate the movement, and then use RSS variance to perform tracking, specifically using our hybrid NN-HMM method. We demonstrate stability over the time of 1 month during which furniture were moved and APs were relocated, all contributing to changes in the RSS signal paths and mean levels drift over time.

Xu et al. [37] have introduced an algorithm called sequential counting or parallel localizing which is shown to simultaneously track multiple subjects. They have used 13 transmitters and 9 receivers to perform their tracking and results demonstrate accurate tracking performance for 1, 2, 3, or 4 moving subjects within two different environments. However, results are reported for only same day calibration and testing. This group acknowledges in their paper that a limitation of their algorithm involves long term drift in the RSS signals that can impact the accuracy of their algorithm. They note that they attempted using a camera to periodically take pictures of the environment to get a ground truth for a given subject's location and thereby run autocorrelation. This has not been implemented in their paper.

Kaltiokallio et al. [10, 11] implemented a radio tomography based passive location system and deployed it within the homes of seniors for a week. Their system is similar to others [1, 9, 12, 19, 40] that are accurate at localizing but require a large number of sensors. The localization system described by Kaltiokallio et al. uses 33 APs total and accounts for drift in RSS over time using a fading model; contrary to our approach which simply uses RSS variance and a simple calibration phase for training the NN classifiers.

In the future, we plan to add system capabilities such as determining between sitting vs. standing or estimating walking speed, which is only briefly mentioned in this paper. Classifying and reporting other activities of daily-living such as washing dishes, cooking, or going to the bathroom are of specific interest to monitoring seniors in their home. While a current limitation of our system is that it only accurately tracks location for a single person, future work is planned for multi-person tracking and multi-person discrimination. We also plan to integrate our time-of-flight body-worn tracking system [7] with the device-free tracking system to optimize performance. For example, the body-worn device may be used for semi-supervised calibration or periodic re-calibration without the need to require a person to walk or stand in specific locations as is done with the current calibration protocol. And lastly, we plan to deploy the system to a larger number of homes within the Oregon Center for Aging and Technology (ORCATECH) Life Laboratory [21] and use the system to assess mobility changes in seniors longitudinally over many years.

ACKNOWLEDGEMENTS

The project described was supported by grant no. 1R41AG035400 from the National Institute for Aging and grant no. ETAC-12-239042 from the Alzheimer's Association. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Aging or the National Institute of Health.

REFERENCES

1. Chen, X.; Edelstein, A.; Li, Y.; Coates, M.; Rabbat, M.; Men, A. Sequential Monte Carlo for Simultaneous Passive Device-Free Tracking and Sensor Localization Using Received Signal Strength Measurements; 10th IEEE International Conference on Information Processing in Sensor Networks (IPSN); 2011. p. 342-353.
2. Hagen, S. Financing Long-Term Care for the Elderly. Congress of the US, Congressional Budget Office; 2004.
3. Hayes TL, Abendroth F, Adami A, Pavel M, Zitzelberger TA, Kaye JA. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. Alzheimer's & Dementia. 2008; 4(6): 395–405.
4. Hightower J, Borriello G. A survey and taxonomy of location systems for ubiquitous computing. IEEE computer. 2001; 34(8):57–66.
5. Hoffman C, Rice D, Sung HY. Persons with chronic conditions. Their prevalence and costs. Jama. 1996; 276(18):1473–1479. [PubMed: 8903258]
6. Ivanov Y, Sorokin A, Wren C, Kaur I. Tracking People in Mixed Modality Systems. proceedings of SPIE on Visual Communications and Image Processing (VCIP). 2007:6508L.
7. Jacobs PG, Paul AS, Wan EA. EmbedRF Position Tracking and Mobility Assessment System: A low-power and low-cost system for indoor pedestrian tracking and mobility assessment. proceedings of The 24th International Technical Meeting of The Institute of Navigation (ION GNSS 2011). 2011:1409–1506.
8. Jacobs, PG.; Wan, EW.; Schafermeyer, E.; Adenwala, F.; Paul, AS.; Preiser, N.; Kaye, J. Measuring in-home walking speed using wall-mounted RF transceiver arrays; accepted in Engineering in Medicine and Biology Society (EMBC) 2014, 36th Annual International Conference of the IEEE; 2014.
9. Kaltiokallio O, Bocca M, Patwari N. A Multi-Scale Spatial Model for RSS-based Device-Free Localization. arXiv. 2013 arXiv:1302.5914v1[cs.NI].
10. Kaltiokallio O, Bocca M, Patwari N. Follow@ grandma: Long-Term Device-Free Localization for Residential Monitoring. 37th IEEE International Workshop on Practical Issues in Building Sensor Network Applications. 2012:991–998.
11. Kaltiokallio, O.; Bocca, M.; Patwari, N. Enhancing The Accuracy of Radio Tomographic Imaging Using Channel Diversity; Proceedings of the 9th IEEE International Conference on Mobile Ad hoc and Sensor Systems (MASS); 2012. p. 254-262.
12. Ke W, Liu G, Fu T. Robust-Sparsity-Based Device-Free Passive Localization in Wireless Networks. Progress in Electromagnetics Research. 2014; C 46:63–73.
13. Kivimäki T, Vuorela T, Peltola P, Vanhala J. A Review on Device-Free Passive Indoor Positioning Methods. International Journal of Smart Home. 2014; 8(1):71–94.
14. Kolodziej, KW.; Hjelm, J. Local positioning systems: LBS applications and services. CRC press; 2010.
15. Kosba, AE.; Saeed, A.; Youssef, M. RASID: A Robust WLAN Device-Free Passive Motion Detection System; 2012 IEEE International Conference on Pervasive Computing and Communications; 2012. p. 180-189.
16. Kosba, AE.; Abdelkader, A.; Youssef, M. Analysis of a device-free passive tracking system in typical wireless environments; 3rd International Conference on New Technologies, Mobility and Security; 2009. p. 1-5.
17. Lakdawalla D, Goldman DP, Bhattacharya J, Hurd MD, Joyce GF, Panis CW. Forecasting the nursing home population. Medical Care. 2003; 41(1):8–20. [PubMed: 12544538]
18. Moussa, M.; Youssef, M. Smart devices for smart environments: Device-free passive detection in real environments; 7th Annual IEEE International Conference on Pervasive Computing and Communication; 2009. p. 1-6.
19. Nannuru, S.; Li, Y.; Coates, M.; Yang, B. Multi-Target Device-free Tracking using Radio Frequency Tomography; Seventh IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP); 2011. p. 508-513.

20. Ohta S, Nakamoto H, Shinagawa Y, Tanikawa T. A health monitoring system for elderly people living alone. Journal of Telemedicine & Telecare. 2002; 8(3):151–156. [PubMed: 12097176]
21. Oregon Center for Aging & Technology (ORCATECH), Oregon Health & Science University, Portland OR. http://www.orcatech.org
22. Patwari N, Wilson J. RF Sensor Networks for Device-Free Localization: Measurements, models, and algorithms. Proceedings of the IEEE. 2010; 98(11):1961–1973.
23. Paul AS, Wan EA, Jacobs PG. Sigma-point Kalman smoothing for indoor tracking and auto-calibration using time-of-flight ranging. In proceedings of The 24th International Technical Meeting of The Institute of Navigation (ION GNSS 2011). 2011:3461–3469.
24. Hagler S, Austin D, Hayes TL, Kaye J, Pavel M. Unobtrusive and Ubiquitous In-Home Monitoring: A Methodology for Continuous Assessment of Gait Velocity in Elders. IEEE Transactions on Biomedical Engineering. 57(4):813–820. [PubMed: 19932989]
25. Sabek I, Youssef M. Spot: An accurate and efficient multi-entity device-free WLAN localization system. *arXiv* preprint arXiv 1207. 2012:4265.
26. Saeed A, Kosba AE, Youssef M. Ichnaea: A Low-Overhead Robust WLAN Device-Free Passive Localization System. IEEE Journal of Selected Topics in Signal Processing. 2014; 8(1):5–15.
27. Savazzi S, Nicoli M, Carminati F, Riva M. A Bayesian Approach to Device-Free Localization: Modeling and Experimental Assessment. IEEE Journal of Selected Topics in Signal Processing. 2014; 8(1):16–29.
28. Scanaill CN, Carew S, Barralon P, Noury N, Lyons N, Lyons GM. A review of approaches to mobility telemonitoring of the elderly in their living environment. Annals of Biomedical Engineering. 2006; 34(4):547–563. [PubMed: 16550450]
29. Seifeldin M, Saeed A, Kosba AE, El-Keyi, Amr, Youssef M. Nuzzer: A Large-Scale Device-Free Passive Localization System for Wireless Environments. IEEE Transactions on Mobile Computing. 2013; 12(7):1321–1334.
30. Sixsmith A. An evaluation of an intelligent home monitoring system. Journal of Telemedicine & Telecare. 2000; 6:63–72. [PubMed: 10824373]
31. Wagner, B.; Timmermann, D.; Ruscher, G.; Kirste, T. Device-Free User Localization Utilizing Artificial Neural Networks and Passive RFID; IEEE Conference on Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS); 2012. p. 1-7.
32. Wan, EA.; Paul, AS. A tag-free solution to unobtrusive indoor tracking using wall-mounted ultrasonic transducers; Proceedings of 2010 IEEE International conference on Indoor Positioning and Indoor Navigation (IPIN); 2010.
33. Wan EA, Paul AS, Jacobs PG. Tag-Free RSSI Based Indoor Localization. proceedings of the 2012 International Technical Meeting of Institute of Navigation. 2012:940–944.
34. Wilson J, Patwari N. A Fade-Level Skew-Laplace Signal Strength Model for Device-Free Localization with Wireless Networks. IEEE Transactions on Mobile Computing. 2012; 11(6):947–958.
35. Wilson J, Patwari N. Radio Tomographic Imaging with Wireless Networks. IEEE Transactions on Mobile Computing. 2009; 9(5):621–632.
36. Woyach K, Puccinelli D, Haenggi M. Sensorless Sensing in Wireless Networks: Implementation and Measurements. 4th IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks. 2006:1–8.
37. Xu, C.; Firner, B.; Moore, RS.; Zhang, Y.; Trappe, W.; Howard, R.; Zhang, F.; An, N. SCPL: Indoor Device-Free Multi-Subject Counting and Localization Using Radio Signal Strength; Proceedings of the 12th international ACM conference on Information Processing in Sensor Networks; 2013. p. 79-90.
38. Yang, Z.; Huang, K.; Wang, G. A New Sparse Reconstruction Algorithm for Device-Free Localization with Sensor Network; 9th IEEE Asian Control Conference (ASCC); 2013. p. 1-6.
39. Youssef, M.; Mah, M.; Agrawala, A. Challenges: Device-free Passive Localization for Wireless Environments; Proceedings of 13th ACM International Conference on Mobile Computing and Networking; 2007. p. 222-229.

40. Zhao, Y.; Patwari, N. Noise Reduction for Variance-Based Device-Free Localization and Tracking; 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON); 2011. p. 179-187.

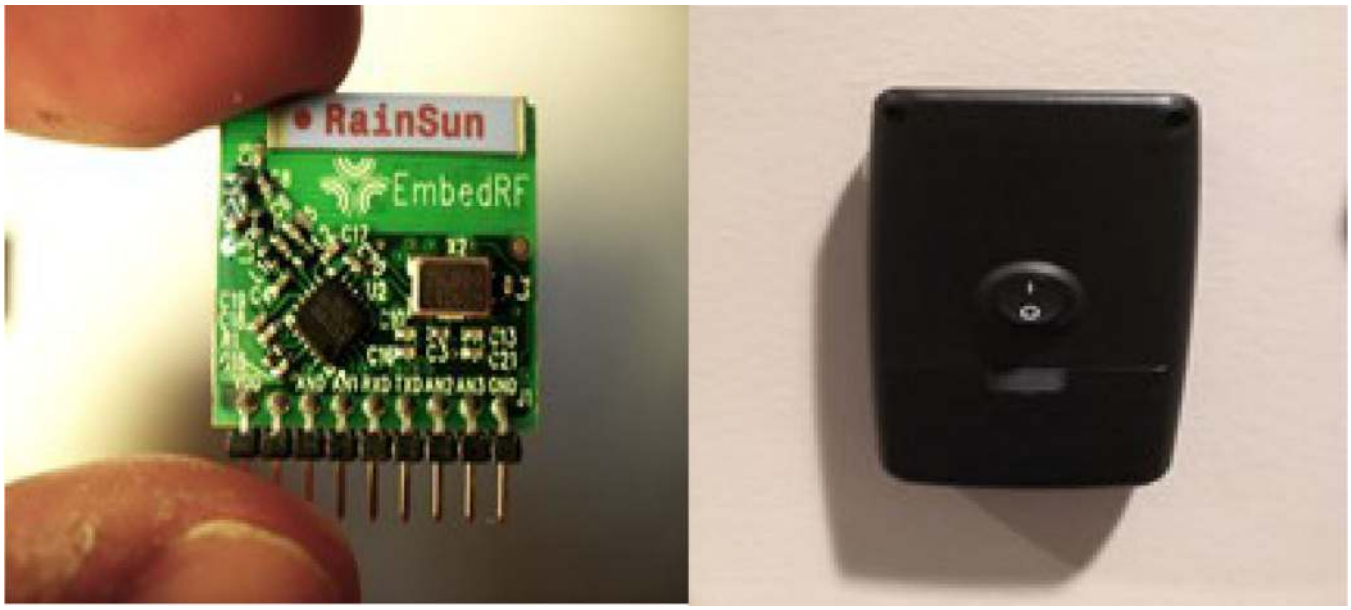

Figure 1.
(a) EmbedRF 915 MHz, wireless transceiver used for the access-points and hub (1.5 grams, 10 byte payload, and 50 ft. range), (b) Enclosure for mounting on the walls.

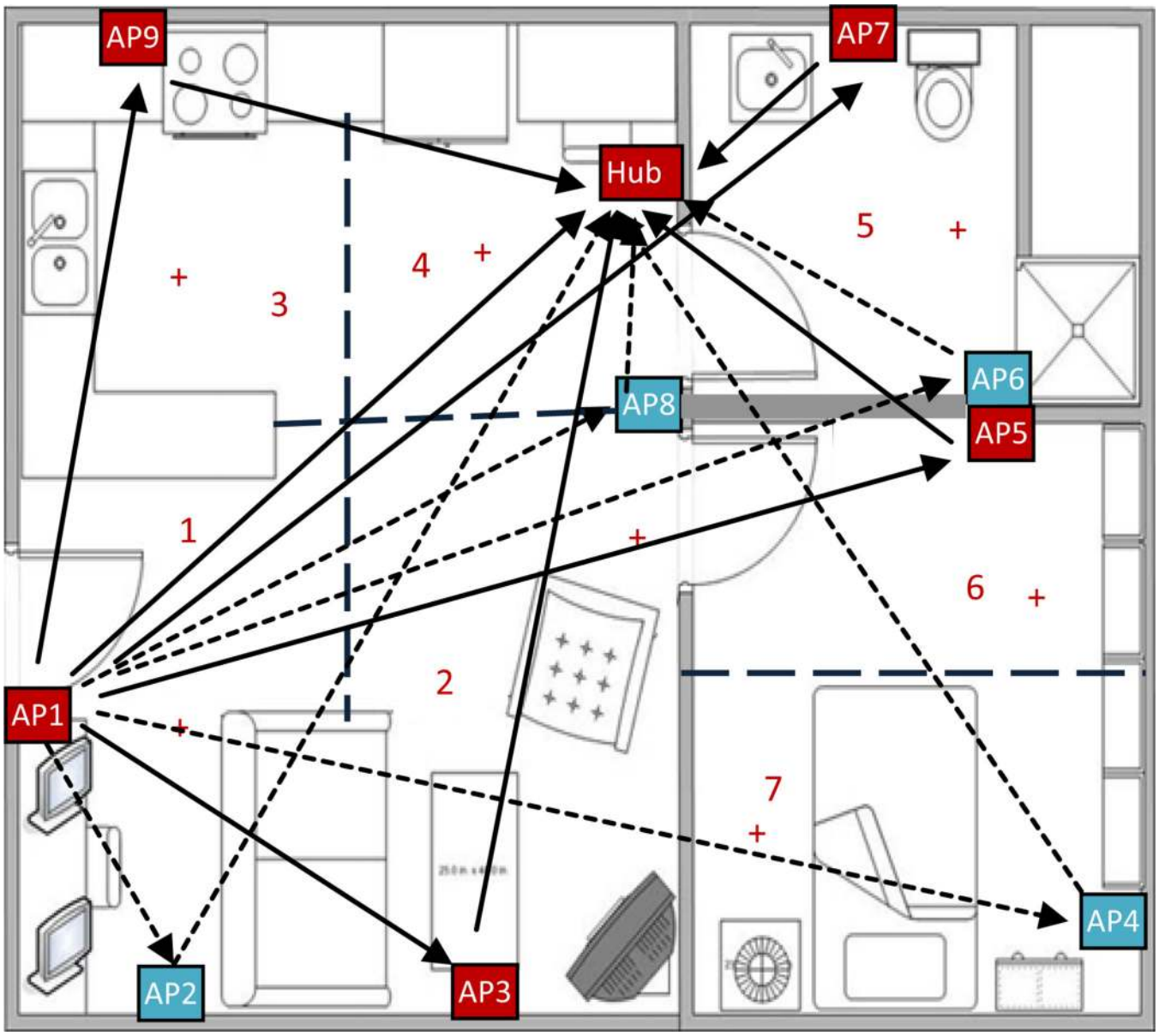

Figure 2.

Room floor plan for the PoCL. Regions are labeled 1–7 in red and are marked with thick dashed lines in the figure. The APs and hub are indicated by red and blue shaded boxes. In some experiments, we used all APs, and in some experiments, we removed the blue APs to see how performance changed with a smaller number of APs. The waypoints 1–7 are labeled with red '+' symbols. The links between transceivers are shown as black lines with arrows or as dashed lines for APs which were removed.

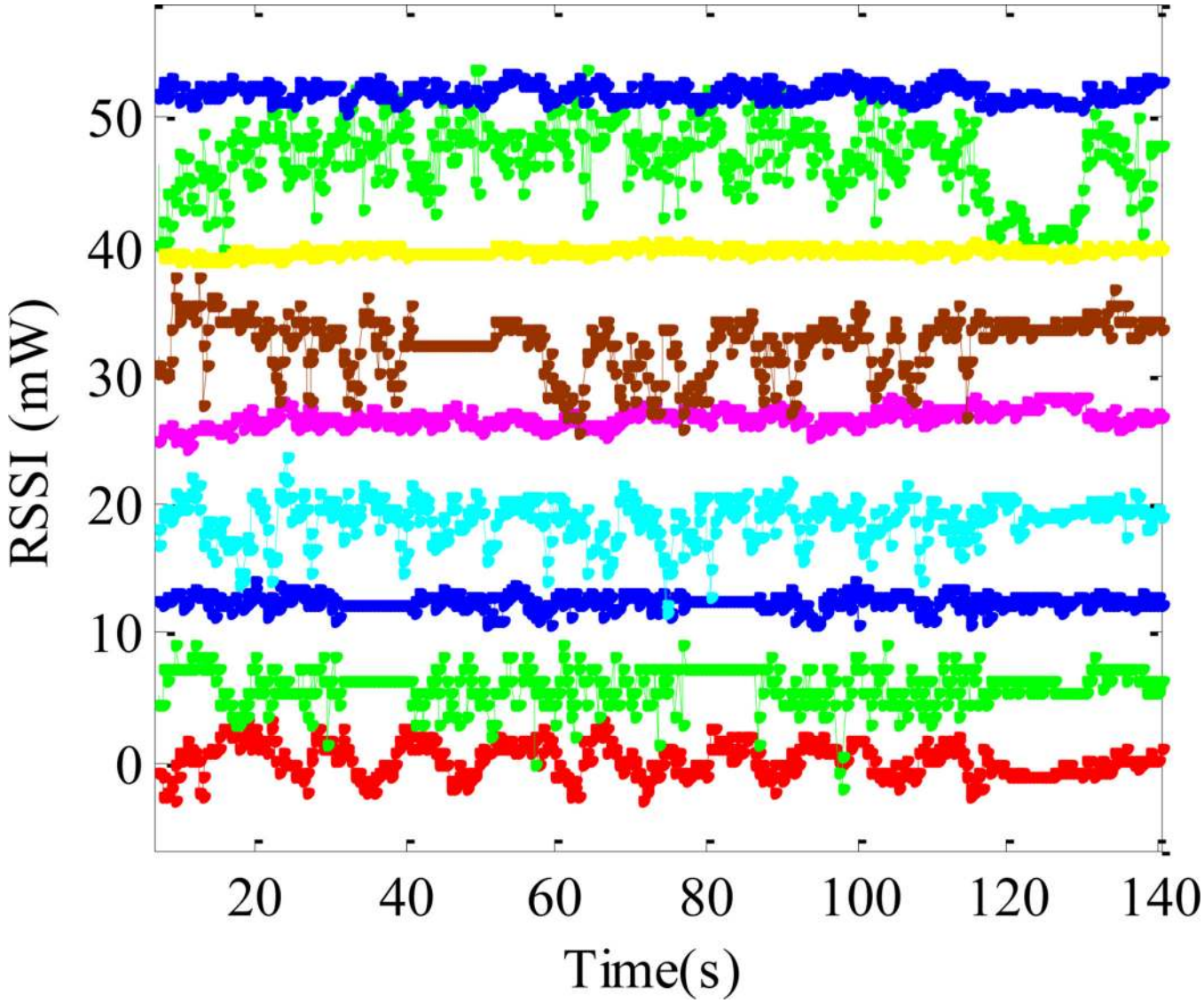

Figure 3.
RSS recordings as a person walks through the PoCL. RSS signals are offset by 5 mW increments in the plot.

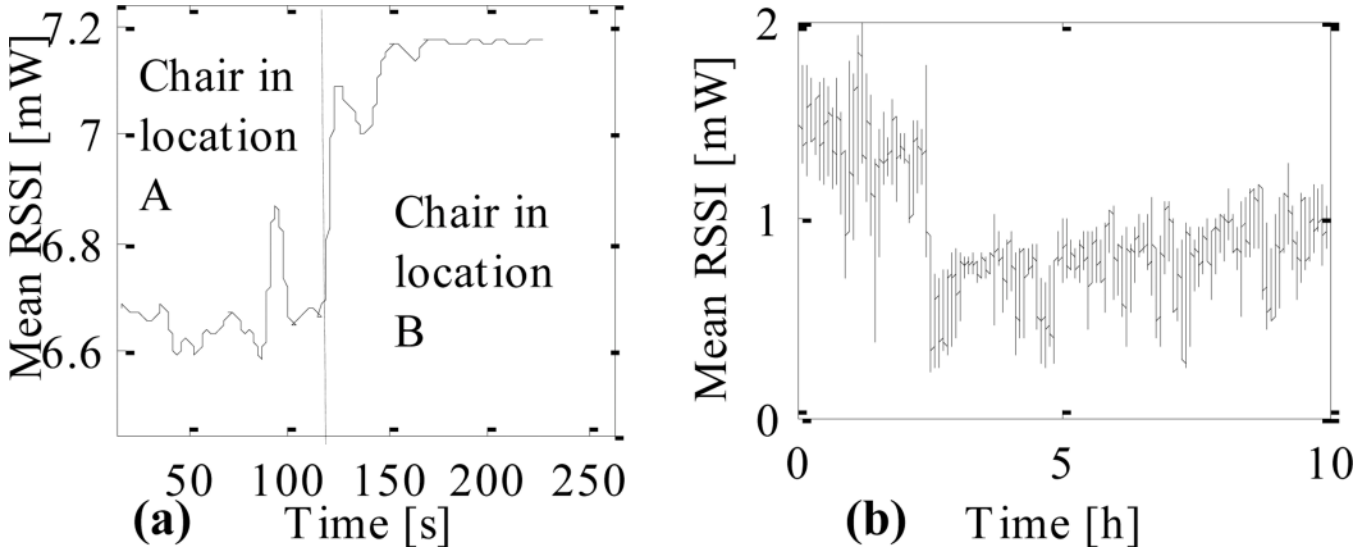

Figure 4.

Example of mean shifts in the RSS (5 sec moving average features plotted). (a) RSS was recorded in an empty room at first with a chair at one location (A), and then with the chair moved to a different location (B). The RSS path shown was not near either location A or B and was therefore likely due to multipath reflections. (b) Long term measurement of a single RSS path for 10 hours. Note the shift in the RSS mean level.

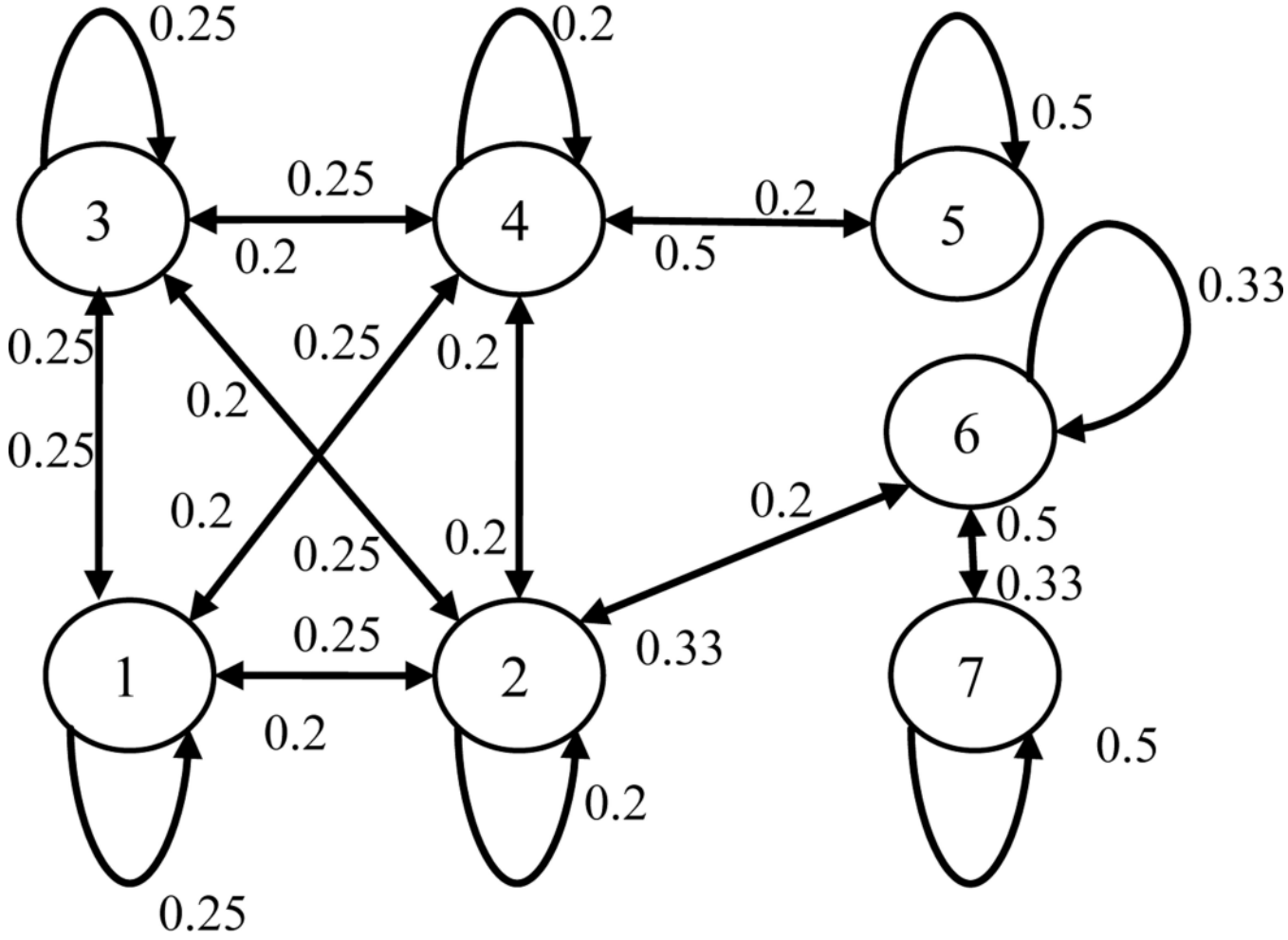

Figure 5.

HMM and transition probabilities for the PoCL room. State transition probabilities from source state to destination state are shown closer to the destination state.

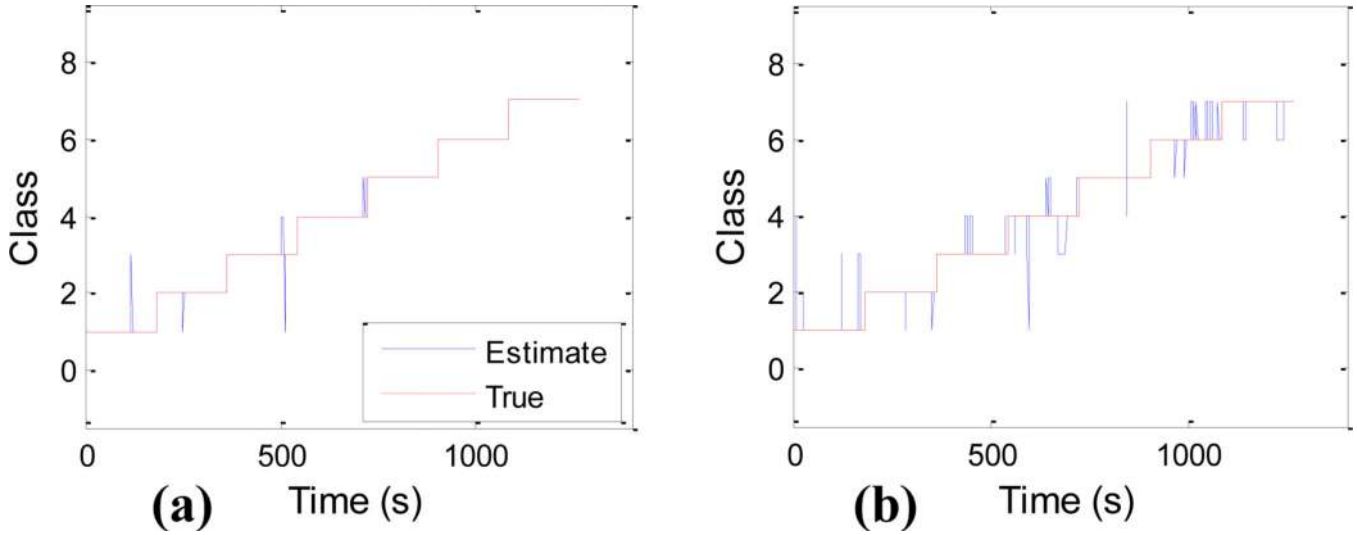

Figure 6.

NN-HMM estimated regions (blue) are shown with ground truth regions (red) during the course of an experiment at PoCL. In this case, the subject moved from region 1–7 sequentially. The subject spent approximately 1 min at each region. (a) Train and test data were collected on the same day (accuracy over all regions is 98.5%), (b) Train and test data were separated by one month (accuracy over all regions is 85%).

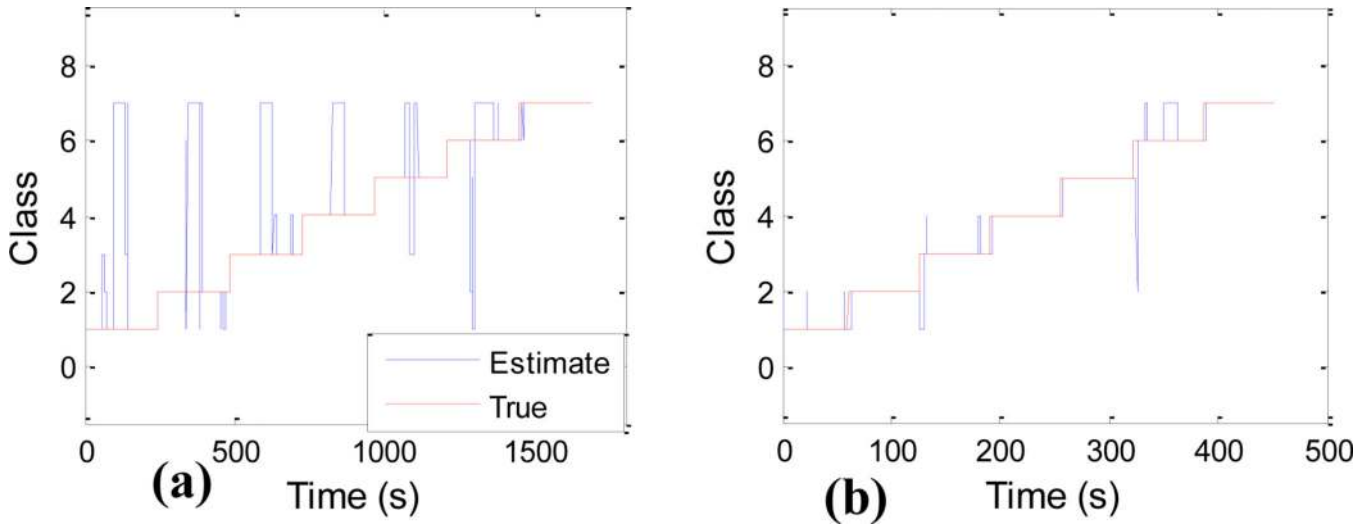

Figure 7.

Classifier estimated regions (blue) are shown with approximate ground truth regions (red) using data obtained at PoCL. The subject moved from regions 1–7 sequentially and performed walking and standing at way-points for approximately 1 min at each region. Train and test data were collected on the same day. Plot (a) shows the accuracy of the standalone NN_r location classifier (76%) and plot (b) displays the hybrid NN-HMM (97.3%). This plot displays the performance improvement of the hybrid NN-HMM.

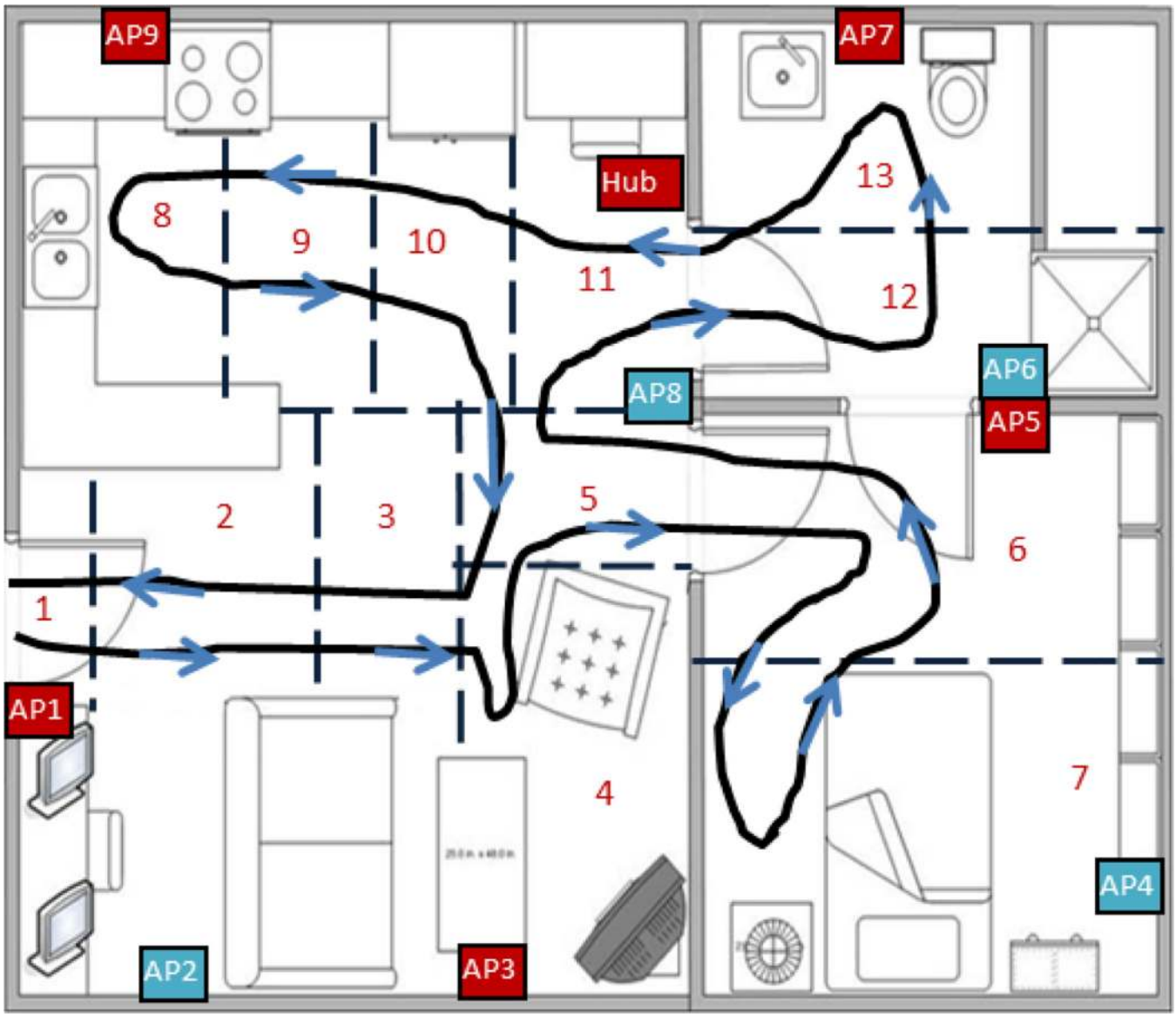

Figure 8.
Approximate walking path in black solid line is overlaid on the PoCL floorplan.

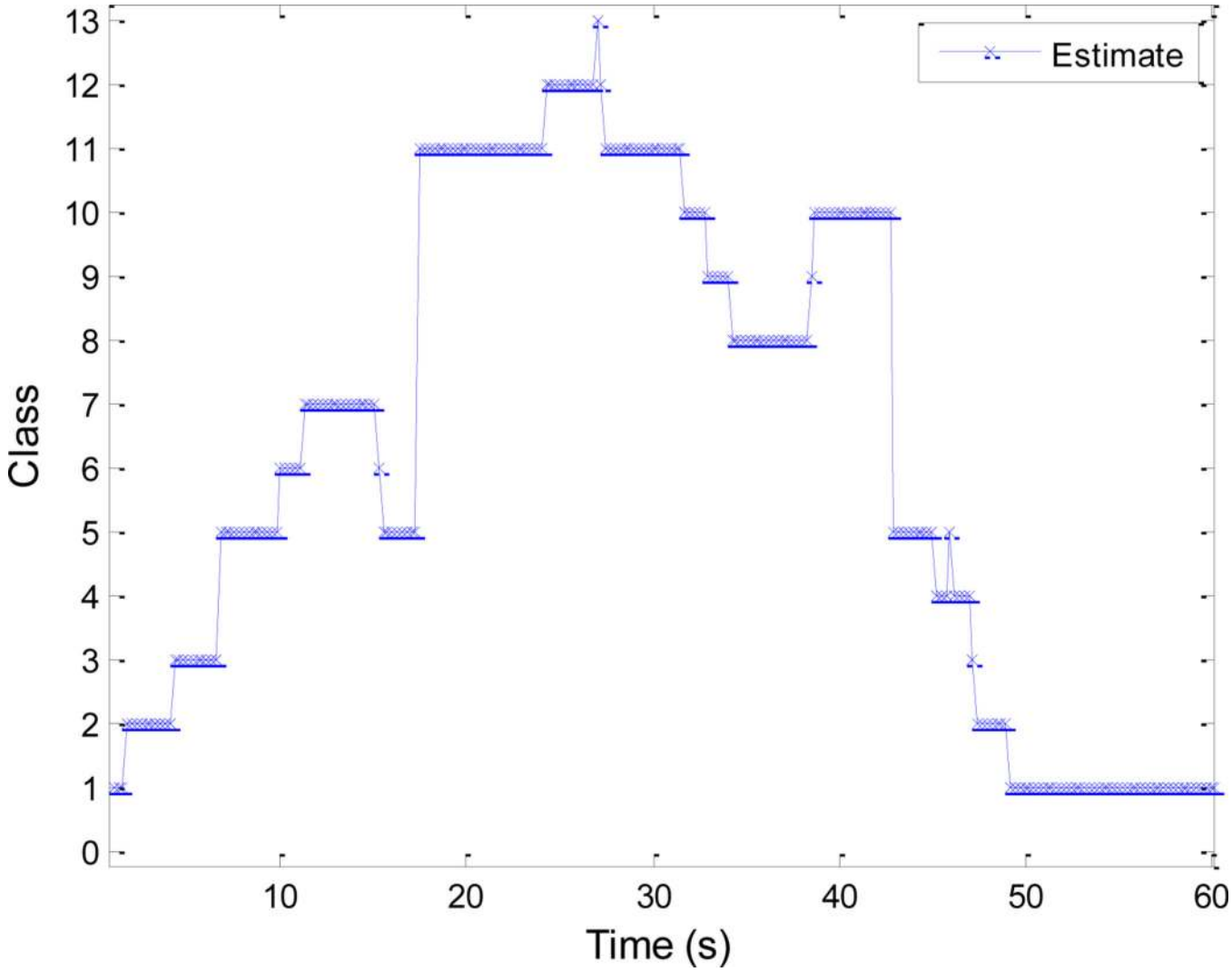

Figure 9.

Classifier estimated regions (blue) are shown. In Figure 8, the true path that the person followed was the following regions:
1-2-3-4-5-6-7-6-5-11-12-13-12-11-10-9-8-9-10-5-4-3-2-1.

Table 1

Confusion matrix in percentage in estimating 0-1-2-3 persons based on experiment performed at PoCL. (E = Estimated, T = True class)

E / T	1 persons	2 persons	3 persons	no person
1 person	93.69	2.25	0.79	3.27
2 persons	8.59	79.01	8.19	4.21
3 persons	3.81	5.83	90.35	0.00
no persons	3.05	0.00	0.00	96.95

Table 2

Confusion matrix in percentage to classify walking vs. standing at PoCL. (E = Estimated, T = True class)

E / T	Walking	Standing
Walking	97.09	2.91
Standing	4.57	95.43

Table 3

Classifier accuracy at PoCL (M=7) for region level localization using short-term variance only (var), mean and variance (mean + var) and mean only (mean) features when testing a) done on the same day, b) separated by 1 week and c) separated by one month. BSPL (M=4) accuracy is only shown using the preferred short-term variance only features.

	a) Same Day	b) 1 Week	c) 1 Month
PoCL (var)	**98.5%**	**95%**	**85%**
PoCL (mean+var)	95%	90%	50%
PoCL (mean)	58%	56%	30%
BSPL (var)	**95.1%**	**91.15%**	**85.5%**

Table 4

Confusion matrix showing classifier accuracy using short-term variance features (in percentage) at PoCL when train and test data were collected on the same day. The true regions are the rows while the estimated regions are the columns. (E = Estimated, T = True region)

E / T	R1	R2	R3	R4	R5	R6	R7
R1	98.90	0.00	1.10	0.00	0.00	0.00	0.00
R2	0.55	99.45	0.00	0.00	0.00	0.00	0.00
R3	0.55	0.00	93.92	5.52	0.00	0.00	0.00
R4	0.00	0.00	0.00	97.24	2.76	0.00	0.00
R5	0.00	0.00	0.00	0.00	100.0	0.00	0.00
R6	0.00	0.00	0.00	0.00	0.00	100.0	0.0
R7	0.00	0.00	0.00	0.00	0.00	0.00	100.0

Table 5

Confusion matrix showing classifier accuracy using short-term variance features (in percentage) at PoCL when train and test data were separated by one month. The true regions are the rows while the estimated regions are the columns. (E = Estimated, T = True region)

E / T	R1	R2	R3	R4	R5	R6	R7
R1	82.87	9.94	5.52	1.66	0.00	0.00	0.00
R2	3.31	96.69	0.00	0.00	0.00	0.00	0.00
R3	0.00	0.00	94.48	5.52	0.00	0.00	0.00
R4	0.55	0.00	14.36	77.35	7.73	0.00	0.00
R5	0.00	0.00	0.00	0.55	98.90	0.00	0.55
R6	0.00	0.00	0.00	0.00	3.31	85.64	11.05
R7	0.00	0.00	0.00	0.00	0.00	13.26	86.74

Proc ACM Int Conf Ubiquitous Comput. Author manuscript; available in PMC 2014 December 24.

Table 6

Comparison of hybrid NN-HMM versus standalone NN_r at PoCL (M=7) in terms of region level localization with testing a) done on the same day, b) separated by one week and c) separated by one month. Performance is also compared when using variance only, mean and variance, and mean only features. Results at BSPL show similar performance.

	a) Same Day	b) 1 Week	c) 1 Month
NN-HMM (var)	**98.5%**	**95%**	**85%**
NN_r(var)	**76%**	**70%**	**62%**
NN-HMM (mean+var)	95%	90%	50%
NN_r (mean+var)	74%	63%	42%
NN-HMM (mean)	58%	56%	30%
NN_r (mean)	45%	40%	30%

Table 7

Classifier performance for region level tracking comparing a) training and testing on the same person, b) train the classifier using 75% less data, 3) training and testing using different people. Train and test data were collected on the same day.

	a) Same Person	b) %75 data	c) Diff. person
BSPL (M=4)	95%	92.2%	91.25%
PoCL (M=7)	98.5%	97%	93.88%