SCPL: Indoor Device-Free Multi-Subject Counting and Localization Using Radio Signal Strength

ABSTRACT

Radio frequency based device-free passive (DfP) localization techniques have shown great potentials in localizing individual human subjects, without requiring them to carry any radio devices. In this study, we extend the DfP technique to count and localize multiple subjects in indoor environments. To address the impact of multipath on indoor radio signals, we adopt a fingerprinting based approach to infer subject locations from observed signal strengths through profiling the environment. When multiple subjects are present, our objective is to use the profiling data collected by *a single* subject to count and localize *multiple* subjects without any extra effort. We propose a successive cancelation based algorithm to iteratively determine the number of subjects. We model indoor human trajectories as a state transition process, exploit indoor human mobility constraints and integrate all information into a conditional random field (CRF) to simultaneously localize multiple subjects. As a result, we call the proposed algorithm *SCPL* – sequential counting, parallel localizing.

We test SCPL with two different indoor settings, one with size 150 m^2 and the other 400 m^2. In each setting, we have four different subjects, walking around in the deployed areas, sometimes with overlapping trajectories. Through extensive experimental results, we show that SCPL can count the present subjects with 86% accuracy when their trajectories are not completely overlapping. Finally, our localization algorithms are also highly accurate, with an average localization error distance of within 1.5 m.

Categories and Subject Descriptors

C.3 [**Special-Purpose and Application-Based Systems**]: Real-time and embedded systems

General Terms

Algorithm, Experimentation, Measurement

Keywords

Device-free, Counting, Passive Localization, Tracking, Multiple Subjects, Fingerprint, Conditional Random Field

1. INTRODUCTION

Ambient Intelligence (AmI) envisions that future smart environments will be sensitive and responsive to the presence of people, thereby enhancing everyday life. Potential applications include eldercare, rescue operations, security enforcement, building occupancy statistics, etc. The key to enable these ubiquitous applications is the ability to localize various subjects and objects in the environment of interest. Device-free passive (DfP) localization [11] has been proposed as a way of detecting and tracking subjects without the need to carry any tags or devices. It has the additional advantage of being unobtrusive while offering good privacy protection.

Several RF-based DfP localization techniques have been proposed [11, 12, 13, 6, 8, 7, 1, 10], and these approaches observe how people disturb the pattern of radio waves in an indoor space and derive their positions accordingly. To do so, they collect training data to profile the deployed area, and form mathematical models to relate observed signal strength values to locations. DfP algorithms can be broadly categorized into two groups: *location-based*, and *link-based*. Location-based DfP schemes collect a radio map with the subject present in various predetermined locations, and then map the test location to one of these trained locations based upon observed radio signals, which is also known as fingerprinting, as studied in [11, 10]. Link-based DfP schemes, however, try to empirically capture the relationship between the radio signal strength (RSS) of a link and whether the subject is on the Line-of-Sight (LoS) of the radio link, and consequently determine the subject's location using geometric approaches [12, 6, 1]. In this study, we adopt the location-based approach because it is more robust to complex multipath environments and thus have better localization accu-

racies. More importantly, this approach is more suitable for tracking multiple subjects.

Many previous studies focus on localizing/tracking a single subject, and tracking multiple subjects has received little attention so far. When we use a location-based DfP scheme to handle multiple subjects, the main challenge lies in the exponential increase in the training overhead if we need to profile the system with different combinations of these subjects. In this study, we propose and evaluate an efficient DfP scheme for tracking multiple subjects using the training data collected by a single subject to avoid expensive training overhead.

Our algorithm consists of two phases. In the first phase, we *count* how many subjects are present using successive cancelation in an iterative fashion. In each iteration, we detect whether the room is empty. If it is not empty, we identify the location for one subject, and then subtract her impact on the RSS values from the collective impact measured in the experiment. Care must be taken when subtracting a subject's impact as the change in the RSS values caused by multiple subjects at the same time is smaller than the sum of RSS changes from each individual subject. In order to compensate for this, we need to multiply a coefficient to a subject's impact and then perform subtraction. The coefficient is specific to the subject's location as well as the link under consideration.

In the second phase, we localize the subjects after their number is known. We partition the deployed area into cells and represent a subject's location using its cell number. We formulate the localization problem as a conditional random field (CRF) by modeling indoor human trajectories as a state transition process and considering mobility constraints such as walls. We then identify the cells occupied by these subjects simultaneously. Since our counting process is sequential and our localization process is parallel, we call our algorithm *SCPL*.

We have tested SCPL in two indoor settings. The first setting is an office environment consisting of cubicles and narrow aisles, which is partitioned into 37 cells. We used the 13 transmitters and 9 receivers that were deployed for some earlier projects. The second setting is an open floor indoor environment, which is partitioned into 56 cells and deployed with 12 transmitters and 8 receivers. In the training phase, we measured the RSS values using a single subject. In the testing phase, we had four subjects with different heights, weights and gender, and implemented four different real life office scenarios. These scenarios all had periods of time when multiple subjects walked side by side and thus had overlapping trajectories. We can count the number of subjects accurately, with a 88% counting percentage when the subjects were not walking side by side, and a 80% counting percentage when they were.

Our localization results have good accuracies, with a mean error distance of 1.3 m considering all the scenarios. We find that it is beneficial to consider indoor human movement constraints according to the floor map when localizing moving subjects and demonstrate 24% improvement on average compared with no floor map information provided.

The rest of the paper is organized as follows. In Section 2, we discuss the applications that benefit from passive localization, our solution framework, as well as related work. Our solution consists of two phases, counting the number of subjects (in Section 3) and localizing the subjects (in Section 4). Then we describe our experimental setup in Section 5 and our detailed results in Section 6. Finally, we provide the concluding remarks and future direction in Section 7.

2. BACKGROUND

Before presenting our SCPL algorithm, we first discuss potential applications, the formulation of the problem and related work.

2.1 Applications that Can Benefit from Passive Localization

Passive localization can find application in many important domains. Below we give a few examples:

Elderly/Health Care: People may fall down in their houses for various reasons, such as tripping, momentary dizziness or overexertion. Without prompt emergency care, this could lead to life-threatening scenarios. Using trajectory based localization information, DfP can perform fall detection quickly because the monitored subject will suddenly disappear from the system.

Indoor Traffic Flow Statistics: Understanding patterns of human indoor movement can be valuable in identifying hot spots and corridors that help energy management and commercial site selection.

Home Security: DfP based home security is a major improvement over camera-based intrusion detection because it can not only detect the intrusion, but also track the intruders, which makes it easier to capture or avoid the intruders.

2.2 Problem Formulation

To solve the passive multi-subject localization problem, we adopt a cell-based fingerprinting approach, similar to the one discussed in [10].

Before we address the multi-subject problem, let us first look at how we localize a single subject. We first partition the deployed area into K cells. In the training phase, we first measure the ambient RSS values for L links when the room is empty. Then a single subject appears in each cell, walks randomly within that cell and takes N RSS measurements from all L radio links. By subtracting the ambient RSS vector from the collected data, we have a profiling dataset $\mathcal{D}$.

$\mathcal{D}$, a $K \times N \times L$ matrix, quantifies how much a single subject impacts the radio RSS values from each cell. Having this profiling dataset $\mathcal{D}$, *we model the subject's presence in cell i as a state S_i* and thus $\mathcal{D} = \{\mathcal{D}_{S_1}, \mathcal{D}_{S_2}, ..., \mathcal{D}_{S_K}\}$. In the testing phase, we first measure the ambient RSS values when the room is empty. Then a subject appears in a random location, and measure the RSS values for all L links while making random moves in that particular cell. Then we subtract the ambient RSS vector from this measured data, and form an RSS vector, O, which shows how much this subject impacts the radio links from this unknown cell. Based on $\mathcal{D}$ and O, we can run classification algorithms to classify the cell number of the unknown cell, thus localizing the subject.

Next we discuss how we extend the same framework to formulate the multi-subject localization problem. In the training phase, our objective is to still *use a single subject's training data* to keep the training overhead low. Taking the training data for different number of subjects will lead to prohibitive overheads, which we will avoid. In the testing phase, multiple subjects appear in random cells, sometimes in the same cell, and we measure the RSS values for all the radio links. We calculate O in the same way as in the single-subject case.

To calculate the locations for these subjects, we need to go through two phases. In the first phase, we identify the number of subjects that are present simultaneously, C, which we call the *counting* phase. In the second phase, we identify in which cells are these C subjects, which we call the *localizing* phase. Please note that subjects are not stationary, but they move around within the deployed area.

2.3 Related Work

Several DfP localization schemes have been proposed and studied in the literature. We, however, point out that SCPL is the first one that counts device-free subjects using radio signal strength, to our best knowledge.

In 2006, Woyach et al. [9] first experimentally demonstrated the feasibility of localizing device-free subjects by observing a difference in RSS changes by a subject moving between (resulting signal shadowing effect) and in the vicinity (causing small-scale fading) of a pair of transmitter and receiver. From then on, several DfP approaches have been proposed in the literature, which can be broadly categorized into two groups as follows.

Link-based DfP: These techniques look for those radio links close to the target subjects and further determine their locations based on the RSS dynamics. Zhang et al. [12] set up a sensor grid array on the ceiling to track subjects on the ground. An "influential" link is one whose RSS variance exceeds a empirical threshold. The authors determine a subject's location based upon the observation that these influential links tend to cluster around the subject. This technique form a consistent link-based model to relate the subject's location relative to the radio link locations, which is referred as to *calibration by link*. Following this intuition, radio tomographic imaging (RTI) [6] has been proposed to use this technique to reconstruct the tomographic image for localizing subjects based on their locations relative to the radio links Line-of-Sight (LoS). RSS attenuation and variance are used as data primitives in [6, 8]. Auxiliary particle filtering technique is leveraged to improve the tracking performance [1]. Calibration by link may suffer from the generalization error when applied in cluttered indoor environments with rich multipath. A typical example is that when a subject randomly block the LoS of three different radio links, the RSS change of those radio links may be negative, or positive, or even be zero, which depend on the composition of the building structure and all the inside objects. In addition, a dense nodes deployment is required to provide enough radio LoS links to cover all the physical space.

Location-Based DfP: This approach is also known as "fingerprinting," a popular approach for RF-based localization. It was first studied in [11] in the context of passive localization. The authors first collect a radio map with the subject present in a few predetermined locations, and then map the test location to one of these trained locations based upon observed radio signals. This method tames multipath effect better than the link-based approaches in cluttered environments because it measures the RSS ground truth for different positions without generalized assumption and thus provide robust accuracies to different indoor settings. In addition, it does not require a nodes deployment as dense as in calibration by link because when the subject is in the position has no intersection with any radio LoS links, the RSS ground truth still can provide a distinguishable record from other positions. In [10], the authors propose a cell-based calibration with random walk method profile the system in order to improve the accuracy and meanwhile reduce the calibration overhead. However, the downside of fingerprinting is also evident: the calibration procedure is tedious, especially for multi-subject in a large environment.

3. COUNTING THE NUMBER OF SUBJECTS

In this section, we first provide empirical data to help the readers understand the impact of having multiple subjects at the same time, and then describe our sequential counting algorithm.

3.1 Understanding the Impact of Multiple Subjects on RSS Values

Let us first understand the relationship between a single subject's impact on the room RSS level and multiple subjects' impact. In particular, we would like to find whether the relationship is linear.

As shown in previous studies such as [11, 12, 6, 1, 10], the RSS level of a radio link changes when a subject is near its

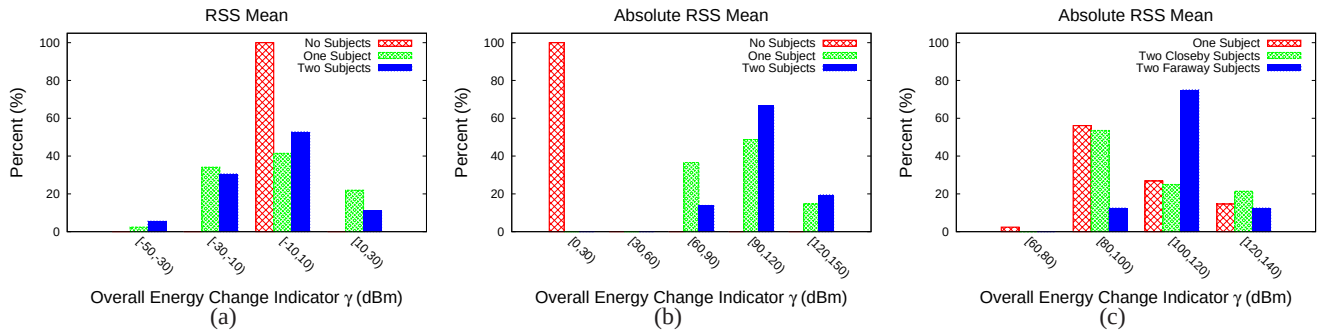


Figure 1: In terms of overall energy change indicator γ, (a) "RSS Mean", for zero, one, and two subjects. (b) "Absolute RSS Mean" for the same measurement shows better discrimination between zero and more than zero subjects. (c) Two subjects separated by more than 4 meters are clearly distinguishable from one subject.

Line-of-Sight (LoS). Based on this observation, we make a simple hypothesis: *more subjects will not only affect a larger number of spatially distributed radio links, but they will also lead to a higher level of RSS change on these links.* If this is true, we can infer the number of subjects that are present from the magnitude of the RSS change that we observe in the deployed area. We use the sum of the individual link RSS change to capture the *total energy change* in the environment as

$$\gamma = \sum_{l=1}^{L} O^l,$$

where O^l is the RSS change on link l.

Next we look at how to capture the RSS change of link l. A straightforward metric is to subtract the mean ambient RSS value for link l (when the room is empty) from the measured mean RSS value for link l, the result of which is referred to as *RSS mean difference*. RSS mean difference is a popular metric that has been used in several studies, e.g., as seen in [11, 6, 1, 10]. However, upon deliberation, we find that RSS mean difference is not suitable for our purpose, mainly because the value is not always positive. Due to the multipath effect, the presence of a subject does not always weaken a link, but sometimes, it may actually strengthen a link! As a result, the RSS mean difference can be negative. In this case, summing up each link's RSS mean difference does not lead to the correct total energy change in the environment because their values may cancel out each other. To address this issue, we thus propose to use *absolute RSS mean difference* which has a more compact data space than RSS mean when a cell is occupied.

Our experimental results confirm that the absolute RSS mean difference is a more suitable metric. In this set of experiments, we collect the RSS values when there are 0, 1 and 2 subjects who make random movements (with pauses) in the deployed area. We compute the corresponding γ value by using both RSS mean difference and absolute RSS mean difference, and plot their histograms in Figures 1(a)-(b) respectively. In Figure 1(a), when the room is empty, we observe γ values $\in [-10, 10)$ which means the overall energy level is rather stable. However, with 40% to 50% of chances, we still observe $\gamma \in [-10, 10)$ when subjects are present. This is because individual RSS mean differences can cancel out each other, and thus their sum is not a good indicator of the total energy change caused by having multiple subjects.

Absolute RSS mean difference is a better metric, as shown in Figure 1(b). The γ value when there are two subjects is statistically greater than the γ value when there is only one subject. As a result, in the rest of this paper, unless explicitly noted, we use absolute RSS mean difference as the metric to capture the RSS change in the environment. Finally, we note that the γ value alone is inadequate to distinguish between one or two subjects.

By looking at the two-subject data more carefully, we can further separate them into two groups based on the distance between the subjects. If the distance is more than 4 meters (we choose this threshold from the data sets), we call the two subjects *faraway*, and call the subjects *nearby* if the distance is less. We then plot the histograms of these groups in Figure 1(c). When subjects are close to each other, more links will be affected by both subjects, and fewer links are affected by only one of the subjects. Consequently, the γ value in this case will be smaller than the γ value when the two subjects are farther apart. Furthermore, we point out that the γ value when we have C subjects at the same time is smaller than the sum of the individual γ value from each subject. As a result, it is hard to distinguish having two subjects close to each other from having only one subject.

In summary, we have two main observations from these experiments. First, the absolute RSS mean difference is a suitable metric to capture the impact caused by the appearance of a subject. Second, the total energy change, γ, reflects the level of impact subjects have in the room, but we cannot rely on the value of γ alone to infer how many subjects are

present.

3.2 Counting Subjects Using Successive Cancelation

We use successive cancelation to count the number of subjects. When multiple subjects coexist, it often so happens that one subject has a stronger influence on the radio signal than the rest. Thus, our counting algorithm goes through several rounds. In each round, we estimate the strongest subject's cell number in this round assuming there is only a single subject, i, and then subtract her RSS contribution to the RSS change from the current RSS vector O to obtain the remaining RSS vector that will be used in the next round.

If this problem were linear, we could simply subtract the mean vector μ_i associated with cell i in the profiling data $\mathcal{D}$ from the observed RSS vector O. However, as shown in the previous subsection, the total impact from multiple subjects is not linear to the number of subjects – the impact observed when C subjects appear at the same time is smaller than the sum of each subject's impact if they appear one at at time. To be more precise, O is an underestimation of the linear combination of the mean values of the associated cells that we collected in $\mathcal{D}$. To address this issue, instead of subtracting μ_i directly from O, we multiply a coefficient that is less than 1 to μ_i and subtract this normalized term from O. This coefficient, however, is not uniform across all the cell and link combinations; instead, it is specific to each cell and link pair because different cells have different impacts on a link. We will then calculate the location-link coefficient matrix, $B = (\beta_{i,l})$ where $\beta_{i,l}$ is the coefficient for cell i and link l.

Our algorithm to calculate the coefficient matrix B is detailed in Algorithm 1. The basic idea is that, for each link l, we compute the correlation between a cell pair, (i, j) with respect to link l. The two cells that both are close to a link are highly correlated with respect to this link. We use h^l_{ij} to denote this correlation[1]. Note that all the RSS values in profiling data are non-negative, and thus we have $h^l_{ij} \geq 0$. For each cell i, we pivot that cell and compute the β_{il} as

$$\beta_{il} = \frac{h^l_{ii}}{\sqrt{\sum_{j=1}^{K} {h^l_{ij}}^2}}.$$

Basically, when two subjects occupy cells i and j respectively, and only one of them affects link l, they have low correlation and the value of h^l_{ij} is close to 0. On the other hand, when they both affect link l, the value of h^l_{ij} will reflect their positive correlation.

Once we determine the location-link coefficient matrix B, we describe our successive cancelation based counting algorithm (shown in Algorithm 2), which can identify the subject count C from the observation RSS vector O using the profiling RSS matrix $\mathcal{D}$ collected by a single subject. We first compute γ^0's and γ^1's from the ambient RSS vector and the profiling RSS matrix $\mathcal{D}$ respectively. Then, we construct a 95% confidence interval for the distribution of γ^0's and γ^1's and refer to the associated lower and upper bounds as c^0_L, c^0_U, c^1_L, c^1_U. From the observation RSS vector, O, we first compute its γ value and then perform a presence detection: if $\gamma \in (c^0_L, c^0_U)$, we claim the room is empty. Otherwise, we will claim there is at least one subject present and start to iteratively count the number of subjects using successive cancelation to finally determine the value of C.

Algorithm 1: Location-Link Correlation Algorithm

input : $\mathcal{D}$- The training data collected from L links among K states/cells
output: B - The location-link coefficient matrix

```
for l = 1 → L do
    h ← zero matrix of K × K
    for i = 1 → K do
        for j = 1 → K do
            I ← training data indices associated with state S_i
            J ← training data indices associated with state S_j
            // Compute the link correlation
            h_ij ← E[D_Il D_Jl]
    for i = 1 → K do
        normfactor ← sqrt( Σ_{j=1}^{K} h_ij^2 )
        // Compute the location-link coefficient for cell i and link l
        β_il ← h_ii / normfactor
```

In each successive cancelation iteration, we do the following:

- *Presence Detection.* We first perform a presence detection by checking if $\gamma \geq c^1_U$ to find out whether there is any more subject in the room. Please note that this condition is stronger than $\gamma \geq c^0_U$, and we will take care of the last iteration separately. If the presence detection returns a 'yes', we increment the detected subject count C, and go to the next step. Otherwise, we end the algorithm.
- *Cell Identification.* If there is a subject in this iteration, we estimate the occupied cell q by

  $$q = \underset{i \in \mathcal{S}}{\operatorname{argmax}} P(O|S_i),$$

  where $\mathcal{S}$ is the set of remaining unoccupied cells.
- *Contribution Subtraction.* Next, we cancel the impact of this subject from cell q by subtracting $\mu_{ql} \cdot \beta_{ql}$ from O^l for each link l.

In the last round, we relax the lower bound of γ^1 by taking the average of c^0_U and c^1_L, which means we consider the possibility that when the last subject is detected in our algorithm,

[1]Notice that we use correlation h^l_{ij} instead of correlation coefficient ρ^l_{ij} because ρ^l_{ii} will always be 1 and thus guarantee its dominance among the all the cells on all the links when the cell i is detected first, which is not true.

the corresponding γ is lower than the c_L^1 but still higher than c_U^0. This further compensates for the over-subtraction in our earlier iterations.

Algorithm 2: Successive Cancelation-Based Device-free Passive Counting Algorithm

input : $\mathcal{D}$- The training data collected from L links among K cells
$\mathcal{S}$- The states $\{S_1, ..., S_K\}$ associated with the K cells
O- The testing data collected from L links when subjects are in unknown locations
B - The estimated location-link coefficient matrix generated from Algorithm 1
c_L^0, c_U^0- The lower and upper bounds of the 95% confidence interval when there is no subjects in the deployed area
c_L^1, c_U^1- The lower and upper bounds of the 95% confidence interval when there is one subject in the deployed area
output: C- The estimated number of subjects present in the deployed area

```
C ← 0
γ ← Σ_{l=1}^{L} O^l
// Presence detection
if γ ∈ (c_L^0, c_U^0) then
    return C;
else if γ ≤ c_L^0 then
    return −1;
// Count the present subjects
else
    while true do
        if γ ≥ c_U^1 then
            // Estimate the most likely occupied cell
            q ← argmax_{i∈S} P(O|S_i)
            // Remove the training data associated with the estimated cell in each round
            D ← D\D_q
            S ← S\q
            // Update the testing data by removing the partial impact caused by the detected subject in each round
            for l = 1 → L do
                O^l ← O^l − β_{ql} μ_{ql}
            C ← C + 1
            // Update the overall affect energy indicator
            γ ← Σ_{l=1}^{L} O^l
        else if γ ∈ ((c_U^0 + c_L^1)/2, c_U^1) then
            C ← C + 1
            return C;
        else
            return −1;
```

4. LOCALIZING MULTIPLE MOVING SUBJECTS WHEN THE SUBJECT COUNT IS KNOWN

In this section, we discuss how we localize multiple moving subjects when the subject count is known. In SCPL, we track multiple subjects in parallel, unlike in the counting phase where we count the number of subjects sequentially. Radio interference is very complex and unpredictable, especially when multiple subjects are present and a link is affected by multiple people. In this case, it is hard to quantify the exact impact of a subject. Even after considering the cell link coefficient matrix B, we may still overestimate (or, underestimate) a subject's impact on a link. These errors, while insignificant enough not to hurt the counting process, will lead to inferior localization results. On the other hand, parallel tracking keeps all the raw RSS values and can provide better results.

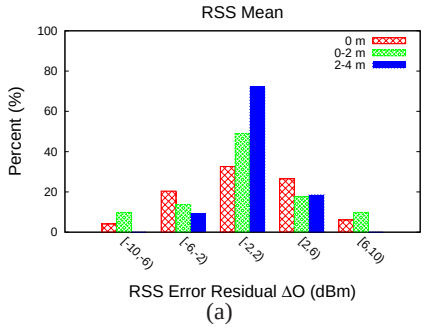
4.1 Understanding the Challenge of Localizing Multiple Subjects

Before presenting our localization algorithm, we first take a closer look at how multiple subjects collectively affect the RSS values and thus complicate the localization problem through empirical data. The complexity of this problem mainly stems from the multi-path effect [4], a typical error source in RF-based indoor localization. In this problem, multi-path can cause nonlinear interference in a radio space when multiple subjects are present. More precisely, when multiple subjects coexist in different locations, the resulting RSS value will not be simply the summation of the individual RSS values from a single subject independently in those locations. The gap between these two is larger when these subjects are close to each other. To validate this conjecture, we randomly select a few positions with certain distances apart. We first have one subject, A, collect the RSS measurements by standing stationary in these locations. Then, we involve another subject, B with similar height and weight as A, and have them stand in two different positions, say i and j. We use O_i and O_j to denote the measured RSS vector when A is standing in positions i and j independently, and O_{ij} the measured RSS vector when A and B are standing in positions i and j simultaneously. In a linear space, vector O_{ij} would be simply the summation of O_i and O_j. However, as mentioned before, this problem is nonlinear, especially when subjects are close to each other. To quantify the degree of nonlinearity, we define the *RSS Error Residual* as

$$\Delta O^l = O_{ij}^l - O_i^l - O_j^l,$$

for link l. A larger ΔO^l value indicates a higher non-linear degree. To articulate the nonlinearity nature, we remove link l if its O_{ij}^l, O_i^l, O_j^l values are all less than 1 because these links are actually not affected by the subjects in any case. We plot the histograms of the remaining O^l values in Figure 2.

From Figure 2, we have three main observations. Firstly, when the two subjects stand side by side (i.e., the distance between them is 0 m), there are only about 30% and 50% chances that we see $|\Delta O^l| < 2$ for RSS mean and absolute RSS mean respectively, which validates our problem is indeed nonlinear. As the distance becomes higher than 2 m, the probability of having $|\Delta O^l| < 2$ rises to more than 70% for both RSS mean difference and absolute RSS mean difference. Secondly, the error residual can be negative under RSS



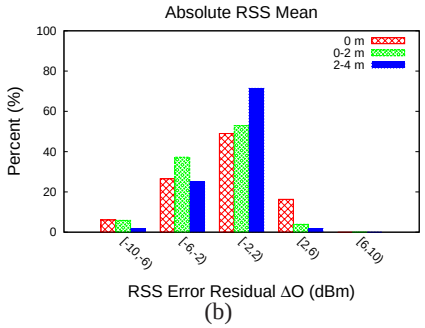


Figure 2: The RSS residual error forms a double-sided distribution when using RSS mean, while it is approximately single-sided distributed using absolute RSS mean.

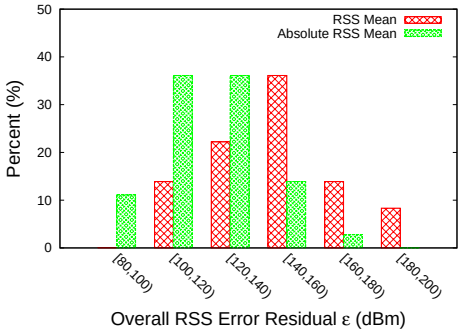


Figure 3: Absolute RSS mean has a smaller overall RSS error residual distribution.

mean difference, but is positive under absolute RSS mean difference in most cases, suggesting O_{ij} is consistently an underestimation of $O_i + O_j$. This property is desirable because it ensures monotoneity.

Finally, we define the *total RSS Error Residual* as:

$$\varepsilon = \sum_{l=1}^{L} |\Delta O^l|,$$

which measures the deviation between the profiling data and the RSS measurement in a multi-subject problem. We plot the histogram in Figure 3 and observe that the absolute RSS mean has a smaller ε value, and thus more appropriate for our purposes.

4.2 Conditional Random Field Formulation

Tracking moving subjects actually introduces new optimization opportunities - we can improve our localization results by considering the fact that human locations from adjacent time intervals should form a continuous trajectory, which can be further modeled as a state transition process under conditional random field (CRF) [3]. CRFs are a type of discriminative undirected probabilistic graphical model. We use them to decode the sequential RSS observations into continuous mobility trajectories.

The first step towards formulating a conditional random field is to form the sensor model and transition model respectively. In our problem, we have K states: $\mathcal{S} = \{S_1, S_2, ..., S_K\}$. In a single-subject problem, state S_i means the subject is located in cell i. The sensor model essentially infers the current state based on the observation RSS vector O. We would like to maximize the likelihood $P(q = S_i|O, \mathcal{D})$ when cell i is occupied. In other words, when the subject is located in cell i in the testing phase, we would like to maximize the probability that the estimated state/cell q matches the actually occupied cell i. We assume the observed RSS vectors in each state follow a multivariate Gaussian with shared covariance, as in [10], and denote

$$\delta_i(O) = P(O|S_i),$$

where

$$P(O|S_i) \sim \mathcal{N}(\mu_i, \Sigma).$$

However, the sensor model is imperfect because of the deep fading effect that can cause estimation error through only a few links[2]. Therefore, the cell associated with the maximum probability might be far from the ground truth.

Next, we look at the transition model. In each clock tick $t = 1, 2, ..., T$, the system makes a transition to state q_t. This process models the movement of a subject – the subject moves to a new cell in each tick. We choose a first order CRF, which means the next cell number depends on the current cell number, rather than any earlier history because we do not want to assume any specific human movement trajectories. In our model, subjects can either walk along a straight line, take turns or wander back and forth.

The subject's trajectory can thus be characterized as a parametric Markov random process with the *transition model* defined as the probability of a transition from state i at time $t-1$ to state j at time t in form of

$$T = P(q_t|q_{t-1}),$$

where

$$T_{ij} = P(q_t = S_j|q_{t-1} = S_i).$$

The intuition here is that people cannot walk through walls or cross rooms in a single tick. We believe these mobility constraints can be used to fix most of the errors in the sensor model caused by deep fades.

In our cell-based approach, we define the following:

Cell neighbors are a list of adjacent cells which can be entered from the current cell without violating mobility constraints.

[2]Because of deep fading from multipath, adjacent points can have dramatically different RSS values, leading to large estimation errors.

Order of neighbor is defined as the number of cells a person must pass through to reach a specific cell from the current cell without violating mobility constraints. We assume the subject moves to a new cell every clock tick. For example, as far as cell i is concerned, the 1-order neighbors include its immediate adjacent cells, and its 2-order neighbors include the immediate adjacent cells of its 1-order neighbors (excluding i and i's first order neighbors).

Cell ring with radius r is defined as the area consisting of i's 1-order neighbors, 2-order neighbors, ..., up to its r-order neighbors.

Let $\Omega_r(i)$ be the cells included in i's r-ring and let $N_r(i)$ be the size of $\Omega_r(i)$. Our transition model thus becomes:

$$T_{ij} = \begin{cases} \frac{1}{N_r(i)} & \text{for} \quad j \in \Omega_r(i) \\ 0 & \text{for} \quad j \notin \Omega_r(i) \end{cases}$$

4.3 Localization Algorithm

Having constructed the sensor model and transition model, we can translate the problem of subject tracking to the problem of finding the most likely sequence of state transitions in a continuous time stream. The *Viterbi algorithm* [2] defines $V_j(t)$, the highest probability of a single path of length t which accounts for the first t observations and ends in state S_j:

$$V_j(t) = \underset{q_1,q_2,\ldots,q_{t-1}}{\operatorname{argmax}} \; P(q_1 q_2 ... q_t = j, O_1 O_2 ... O_t | T, \delta).$$

By induction

$$V_j(1) = \delta_j(O_1),$$

$$V_j(t+1) = \underset{i}{\operatorname{argmax}} \, V_i(t) T_{ij} \delta_j(O_{t+1}).$$

Generalizing to the multi-subject case, we denote $\delta_{1:K}(O) = \{\delta_1(O), \delta_2(O), ..., \delta_K(O)\}$ from the sensor model to represent the likelihood of each state. We denote $Q = \{q^1, ..., q^C\}$, where C is the total number of present subjects. For the current state Q_t, we have $\binom{K}{C}$ possible permutations of subject locations. For each permutation j, we denote $Q_j = \{q^1, ..., q^C\}$ and compute the Viterbi score

$$F_j = \sum_{i=1}^{C} \delta_{q_t^i}(O_t) T_{q_{t-1}^i q_t^i}.$$

We then pick the j value that is associated with the maximum Viterbi score as the current state.

We describe our device-free multi-subject localization algorithm in Algorithm 3. We believe we can achieve best localization results when we consider 1-order ring or 2-order ring, which is also confirmed by our experimental results presented in Section 6.

5. EXPERIMENTAL SETUP

In this section, we briefly describe the experimental setup, the data collection process and the metric we use for performance evaluation.

Algorithm 3: Trajectory-Based Device-free Multi-subject Localization Algorithm

input : $\mathcal{D}$- The training data collected from L links among K cells
$O_{1:t}$- The testing data collected from L links when subjects are in unknown locations
C- The estimated number of present subjects in the deployed area
Q_1- The initial state(s) of the present subjects
output: $Q_{1:t}$- The most like sequence of the trajectories of the present subjects

for $i = 2 \rightarrow t$ **do**
  $\delta_{1:K}(O_i) \leftarrow P(O_i | S_{1:K})$
  $\Pi \leftarrow$ is the set of all the possible permutations of $\binom{K}{C}$
  $Q_i \leftarrow \operatorname{argmax}_{j \in \Pi}$ `ViterbiScore`$(Q_{i-1}, Q_j, \delta_{1:K}(O_i), T)$

5.1 Experimental Setup and Data Collection

The radio devices used in our experiments contain a Chipcon CC1100 radio transceiver and a 16-bit Silicon Laboratories C8051-F321 microprocessor powered by a 20 mm diameter lithium coin cell battery, the CR2032. The receivers have a USB connector for loss-free data collection but are otherwise identical to the transmitters. In our experiments, the radio operates in the unlicensed bands at 909.1 MHz. Transmitters use MSK modulation, a 250kbps data rate, and a programmed output power of 0dBm. Each transmitter periodically broadcasts a 10-byte packet (8 bytes of sync and preamble and 2 bytes of payload consisting of transmitter's id and sequence number) every 100 millisecond. When the receiver receives a packet, it measures the RSS values and wraps the transmitter id, receiver id, RSS, timestamp (on the receiver side) into a "data packet". This packet is sent to the centralized system over direct USB connection or through network hubs for data collection and analysis. In our experiments, the RSS data is collected as a mean value over a 1 second window for each link. We choose a 1 second window because a normal person can at most walk cross one cell during a second. For each cell, we collect 30 RSS vectors as the profiling data.

5.2 Performance Metrics

We use the following performance metrics to measure our counting and localizing algorithms.

Counting Percentage is given by:

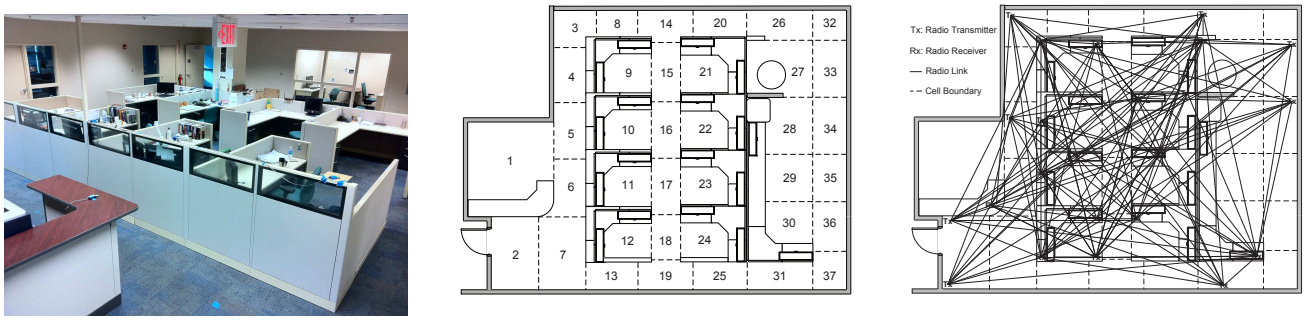
$$1 - \frac{|\hat{C} - C|}{C},$$

where $\hat{C}$ is the estimated subject count and C is the actual subject count.

Error Distance is defined similar to the metric in [5]:

$$d(Q, \hat{Q}) = \frac{1}{C} \min_{\pi \in \Pi} \sum_{i=1}^{C} d(q^i, \hat{q}^{\pi(i)}),$$

where Π includes all the possible permutations of $\{1, 2, ..., C\}$,



(a) Test Field (b) Cell Locations (c) Radio Link Distribution

Figure 4: In (a), we show the office in which we deployed our system. In (b), we show that the office deployment region is partitioned into 37 cubicle-sized cells of interest. In (c), we show the locations of the pre-installed 13 radio transmitters, 9 radio receivers and the corresponding Line-of-Sight links.

$d(q, \hat{q})$ is the Euclidean distance between the ground truth q and the estimated position $\hat{q}$. $Q = \{q^1, q^2, ..., q^C\}$ and $\hat{Q} = \{\hat{q}^1, \hat{q}^2, ..., \hat{q}^C\}$ are within the pre-profiled finite states $\mathcal{S} = \{S_1, S_2, ..., S_K\}$.

6. EXPERIMENTAL RESULTS

In this section, we summarize the results we have obtained from two indoor settings. In each setting, we had multiple subjects each walking along a trajectory.

6.1 Counting and Tracking in an Office Setting

Our first setting is a typical office environment, consisting of cubicles and aisles with a total area of 150 m^2. The environment is quite cluttered as shown in Figure 4(a). The area is broken down to 37 cells such as cubicles and aisle segments, as shown in Figure 4(b). We utilized 13 radio transmitters and 9 radio receivers, whose locations and corresponding link LoS's are shown in Figure 4(c). Here, we need to point out that these devices were installed for some earlier projects, not specifically for this one, and therefore, the link density per cell is non-uniform. This, however, represents a more realistic setting, through which we can show that SCPL can achieve good results without dedicated sensor deployment.

We had four subjects (A, B, C and D) in this series of experiments. We went through several example scenarios and illustrate them in Figure 5:

- *One Subject Scenario:* A left her boss's office, and walked along the aisle to her cubicle.
- *Two Subject Scenario:* When B entered the room, A was walking on the aisle towards him. B waited until they met and walked together for some time, and then separated to go back to their own seats.

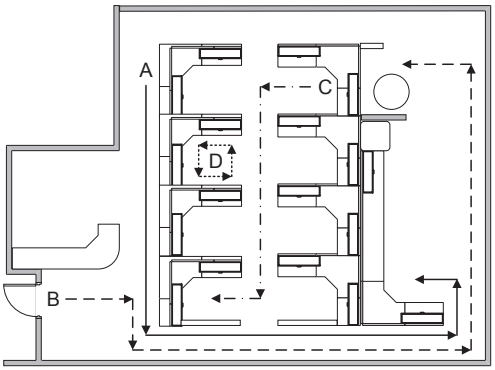


Figure 5: We show the experimental trajectories of subjects A, B, C and D in the office setting. Note the segments of the paths where overlap can occur if the subjects are present at the same time.

- *Three Subject Scenario:* While A and B followed the movement patterns in the above two subject scenario, C walked on the other aisle from one cubicle to another.
- *Four Subject Scenario:* While A, B, and C followed the movement patterns in the above three subject scenario, D was sitting on her seat.

6.1.1 *Counting Results*

The difficulty of subject counting increases when multiple subjects walk together (in the same cell). Thus, we present our counting results in the following three ways: (a) all the experimental data (referred to as *mixed*), (b) the experimental data for when multiple subjects walked together and thus had overlapping trajectories (referred to as *overlap trajectory*), and (c) the experimental data for when multiple subject trajectories did not overlap (referred to as *non-overlap*

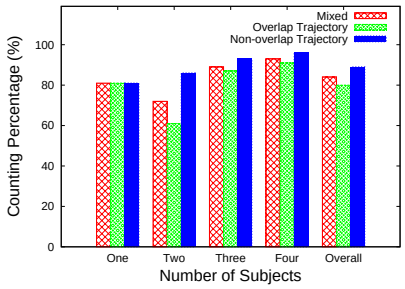


Figure 6: Counting percentage for different number of subjects in the office setting.

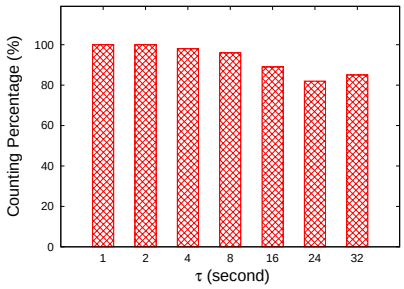


Figure 7: Counting percentage for mixed trajectories using moving window in the office setting.

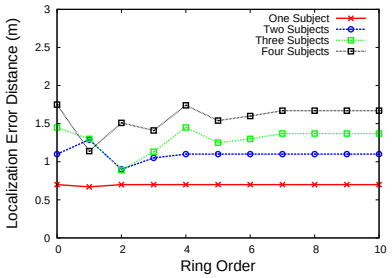


Figure 8: Localization error using different ring order in the office setting.

trajectory). Figure 6 shows the counting percentages in all three cases.

We observe that when we have multiple subjects, the counting percentage is higher in the non-overlap trajectory case. The average counting percentage across all cases is 84%, the average counting percentage for non-overlap cases is 89%, and the average counting percentage for overlap cases is 80%.

Next, we show how the average counting percentage changes as movement patterns change in time. Our four-subject experiment last 32 seconds, and Figure 7 shows the overall average counting percentage (in the mixed case) in the first 1, 2, 4, 8, 16, 24, and 32 seconds. We find that the average counting percentage is very good (until the first 16 seconds), and then drops in the second half of the experiments. This is because two subjects started to merge their trajectories a little while ago. From Figure 7, we note that in an experiment, unless the subjects are always together, we can count the number of subjects accurately during the times when they are apart. The errors caused by temporarily clustered subjects can also be easily addressed. We should run the counting algorithm every second. When we notice the number of subjects suddenly drops, we check their locations before the sudden drop. If no subject's location was close to the exit, then we can conclude that two or more (depending upon the change in the count) joined each other and started walking together. Of course, this information should be validated from the subject location information.

6.1.2 Localization Results

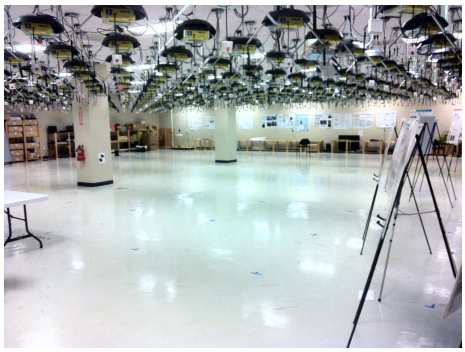
We show the mean of localization error distances in Figure 8 with different ring order parameters. In our setting, we choose 10 as the upper bound of the ring order because all cells are within 10 hops of each other.

Our first observation is that the use of the trajectory information can improve the localization performance by 13.6% – the overall mean localization error distance drops from $1.25m$ (with 0-order ring) to about $1.08m$ (with 1-order ring). We note that the error distance for a single subject does not benefit from using trajectory information because the profiling data is good enough for this case [11, 10]. Multiple subjects, especially when they are close to each other, will cause non-linear radio interference, and thus the data collected from the mutually affected links alone cannot give very accurate localization results. Therefore, the sensor model alone is insufficient for high accuracies. Secondly, we observe that the localization results are less accurate in those cells with lower radio link densities, such as in cell 34-37, because subjects may cause negligible changes to the RSS space at a few points in those cells. Thirdly, trajectory information help prevent the error distance increases dramatically as the increasing number of subjects. Finally, our environment is an office space consisting of cubicles and aisles, and the possible paths a subject can take are rather limited. As a result, we achieve the best localization accuracies with ring order of 1, 2 or 3. Due to the movement constraints, a higher ring order has the same result as not considering any neighbors at all (i.e., 0 ring order). We hypothesize that this may not be true in a more open indoor environment such as (large) homes, malls and museums.

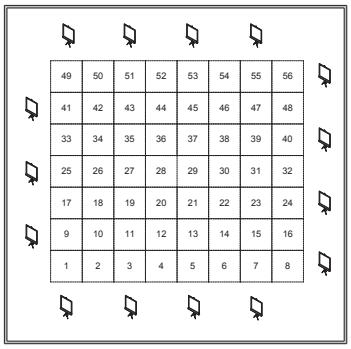
6.2 Counting and Tracking In Open Floor Space

The second test setting is a more open floor of total 400 m^2, as shown in Figure 9(a). We used this setting to model an open hall with a few posters on exhibition, and SCPL can be used to detect traffic flow and infer the most popular poster. The space was partitioned into a uniform grid of 56 cells, as shown in Figure 9(b). We deployed 12 transmitters and 8 receivers in such a way that the link density has a relatively even distribution across the cells, as shown in Figure 9(c). We would like to point out that we used fewer devices in this setting than in the previous one, though this one had a larger area.

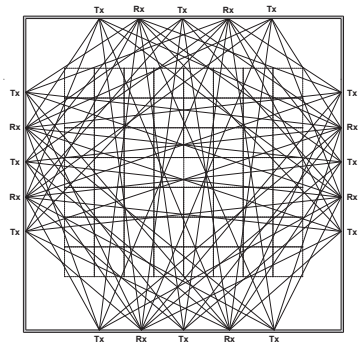
We involved four different subjects in this test and their trajectories are shown in Figure 10. We repeated the same four scenarios as in the previous setting. We plot our counting results in Figure 11. We achieve a 100% counting percentage when there was only a single subject, which is better than the previous setting because the link density is more even in this case. We achieve a counting percentage of 83%, 81%, and 82% for two, three and four subjects respectively,



(a) Test Field



(b) Cell Locations



(c) Radio Link Distribution

Figure 9: In (a), we show the open floor space used for poster exhibition in which we deployed our system. In (b), we show that the deployment region is partitioned into a uniform grid of 56 cells. In (c), we show the locations of the 12 radio transmitters, 8 radio receivers and the corresponding Line-of-Sight links.

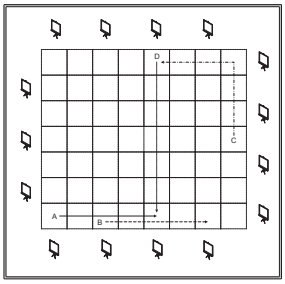


Figure 10: We show the experimental trajectories of subjects A, B, C and D in the open floor space.

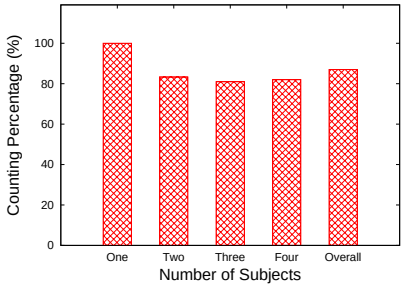


Figure 11: Counting percentage for different number of subjects in the open floor space.

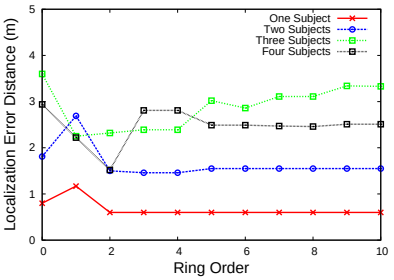


Figure 12: Localization error using different ring order in the open floor space.

resulting in a 87% counting percentage in total.

We present the localization results in Figure 12. In the localization part, we observe similar patterns as in the previous setting: we achieve better localization accuracy using trajectory information. We achieve the best localization accuracy when we adopt the 2-order ring, which is 1.49 m, a 35% improved compared to the 0-ring case.

7. CONCLUSION

In this paper, we present SCPL, an accurate counting and localization system for device-free subjects. We demonstrate the feasibility of using the profiling data collected with only a single subject present to count and localize multiple subjects in the same environment with no extra hardware or data collection. Through extensive experimental results, we show that SCPL works well in two different typical indoor environments of 150 m^2 (office cubicles) and 400 m^2 (open floor plan) deployed using an infrastructure of only 20 to 22 devices. In both spaces, we can achieve about an 86% average counting percentage and 1.3 m average localization error distance for up to 4 subjects. Finally, we shows that though a complex environment like the office cubicles is expected to have worse radio propagation, we can leverage the increased mobility constraints that go with a complex environment to maintain or even improve accuracy in these situations.

Finally, we point out that if we rely on a single subject's training data, the number of subjects that can be accurately counted and localized is rather limited. We had success with up to 4 subjects, but were not very successful with more subjects. In our future work, we will look at how we can accurately localize a larger number of subjects with reasonable overheads.

8. REFERENCES

[1] X. Chen, A. Edelstein, Y. Li, M. Coates, M. Rabbat, and A. Men. Sequential monte carlo for simultaneous passive device-free tracking and sensor localization using received signal strength measurements. In *Proceedings of the 10th international conference on Information Processing in Sensor Networks (IPSN)*, pages 342 –353, April 2011.

[2] J. Forney, G.D. The viterbi algorithm. *Proceedings of the IEEE*, 61(3):268 – 278, march 1973.

[3] J. D. Lafferty, A. McCallum, and F. C. N. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01, pages 282–289, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

[4] T. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2nd edition, 2001.

[5] D. Schuhmacher, B.-T. Vo, and B.-N. Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE Transactions on Signal Processing*, 56(8):3447 –3457, aug. 2008.

[6] J. Wilson and N. Patwari. Radio tomographic imaging with wireless networks. *IEEE Transactions on Mobile Computing*, 9(5):621–632, May 2010.

[7] J. Wilson and N. Patwari. A fade level skew-laplace signal strength model for device-free localization with wireless networks. *IEEE Transactions on Mobile Computing*, PP(99):1, 2011.

[8] J. Wilson and N. Patwari. See-through walls: Motion tracking using variance-based radio tomography networks. *IEEE Transactions on Mobile Computing*, 10(5):612 –621, may 2011.

[9] K. Woyach, D. Puccinelli, and M. Haenggi. Sensorless sensing in wireless networks: Implementation and measurements. In *4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2006*, pages 1 – 8, 2006.

[10] C. Xu, B. Firner, Y. Zhang, R. Howard, J. Li, and X. Lin. Improving rf-based device-free passive localization in cluttered indoor environments through probabilistic classification methods. In *Proceedings of the 11th international conference on Information Processing in Sensor Networks*, IPSN '12, pages 209–220, New York, NY, USA, 2012. ACM.

[11] M. Youssef, M. Mah, and A. Agrawala. Challenges: device-free passive localization for wireless environments. In *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, MobiCom '07, pages 222–229, New York, NY, USA, 2007. ACM.

[12] D. Zhang, J. Ma, Q. Chen, and L. M. Ni. An rf-based system for tracking transceiver-free objects. In *Fifth Annual IEEE International Conference on Pervasive Computing and Communications, 2007. PerCom '07.*, pages 135 –144, march 2007.

[13] D. Zhang and L. M. Ni. Dynamic clustering for tracking multiple transceiver-free objects. In *Proceedings of the 2009 IEEE International Conference on Pervasive Computing and Communications*, pages 1–8, Washington, DC, USA, 2009. IEEE Computer Society.