A Geometric Approach to Device-Free Motion Localization Using Signal Strength

Robert S. Moore, Richard Howard, Pavel Kuksa, Richard P. Martin

Rutgers University Department of Computer Science
Technical Report DCS-TR-674
September, 2010

Abstract:

In this work we describe and evaluate an approach to accurately infer the position in a building where human motion occurs. Our approach does not require the humans to wear any type of device. Such passive mobility localization is applicable in a wide variety of application domains, including those in security, human workflows, and systems management. We position human motion using the change in standard deviation of the received signal strength between stationary transmitters and receivers at known locations. Using a modest transmission rate of once per second, we localize the motion at 2-5 second timescales using a lines-intersecting-tiles method where each line is a straight path between a transmitter and receiver. Our algorithm returns a set of rectangular tiles where the motion has occurred. We experimentally validate our scheme in two different building environments, one containing a cluttered space and a second with a more open arrangement. We show good results for basic mobility detection, with a low number of false positives and negatives. We show that we can localize human motion with a median error of less than 20 ft. We can achieve these results with a modest density of inexpensive active RFID tags, one per 500 ft.^{2}. We also explored how our results degrade with reduced density of transmitters and receivers, and show our mobility detection rates remain good although the geometric precision of the results degrades in line with the density of transmitters.

1 Introduction

Passive motion detection, that is, detecting mobility of objects with no attached wireless devices, and the localization of such motion, are key enablers of applications in areas including security, inventory control, energy management, and human workflow analysis. For example, an energy management application might need the long term human movement patterns in order to make scheduling decisions about which parts of building to heat and cool. A health care workflow application may need to know which rooms are occupied. A data center application may route work away from regions of a building where it detects human motion in order to reduce the chance of an operator-induced fault. In all these domains, localizing human mobility would give developers a new and powerful set of tools towards realizing these emerging applications.

In this work we explore the use of inexpensive commercial off the shelf (COTS) wireless technologies to passively estimate the location of human motion in indoor environments. Our approach centers around first detecting motion, and then returning an area constraining the location. We are able to perform motion detection and localization in real-time, using no offline training, a relatively low density of tags, and generate results with modest geometric error in location estimation.

Often, both cost and privacy motivate using a wireless approach. Other motion detection modalities, such as cameras, ultrasound, IR, and lasers, are limited by line-of-sight (LOS). In many building environments, the LOS limitation thus requires deployment of a large number of sensors. A second issue with cameras is that the potential for misuse and the resulting privacy violations are often too high a risk for many users. Our approach estimates the positions where mobility occurs by observing changing radio signals caused by human motion. Because the frequencies used in COTS wireless devices pass through common construction materials, our approach does not suffer from LOS limitations. In addition, by utilizing technologies that do not necessitate recording personal or private information about the users, we avoid many privacy issues raised by indoor tracking systems that utilize cameras or other sensors that capture user information.

Signal propagation theory predicts mobility detection should be an easier problem than static localization, and this work demonstrates systems built with mobility as the primary detection event provide near perfect recall and precision, as well as obtain good localization perfor-

mance in multiple environments with zero training. Our approach is to deploy low cost active RFID tags and readers at known locations, and then reason about mobility from changes in the standard deviation (σ) of the signal strength. Our geometric approach then provides a set of rectangular tiles where mobility has likely occurred from a global set of overlapping tiles spanning the entire space being monitored. We use a very simple approximation of signal propagation: the highest-weighted regions containing mobility are computed from the weighted sums of the straight line paths between transmitters and receivers over the tile. Although propagation theory predicts more elliptical shapes [6], we found straight line paths an acceptable approximation, and these also greatly simplify the algorithm so it is realizable in real-time.

This enables a large spatial coverage with a less expensive infrastructure. The trade off of this approach is in the accuracy and precision of the motion estimation, though in this work we show we can achieve near perfect motion detection and localize the motion to within 20 ft., which is sufficient for many applications.

We perform direct experiments on a deployed system that provides real-time results, as opposed to using a trace-driven algorithmic approach or simulation. We show results in two different office buildings, one with an average tag density 1 per 250 ft.^{2}, and a second with a tag every 500 ft.^{2}. The base station density was also relatively modest, about 1/550 ft.^{2} We also show our results with a reduced tag densities of 1 per 750 ft.^{2}. The two buildings used as experimental spaces were office environments, one approximately 17,000 ft.^{2} and the other approximately 12,000 ft.^{2}. We performed experiments involving 1, 2 and 3 persons actively walking through the areas and attempting to detect and localize the actors.

Because we used a live system in buildings actively used for research and business, our layout was irregular, as we did not have access to all areas. We found that at an average density of 1 transmitter per 750 ft.^{2} a single 'blind spot' significantly impacted the results. Although the blind spots could be addressed with a careful layout, this would require much more post-deployment testing. We also found the geometry of detection improved as both the density increases and the latency increases (*i.e.* allowing events to be detected after a few seconds), although the gains were somewhat modest as the density doubled from 1 tag/500 ft.^{2} to 1/250 ft.^{2}.

We found that the detection latency was on the order of a few seconds. Even with the modest beacon interval of 1/s, we can detect relatively short motion events with latencies of 2-5 seconds. We anticipated that latency would have a significant effect on the performance of the system due to the 1 second beacon frequency and other network or hardware-related delays. A key result of using direct experiments is that we validated these low latencies, as well as show that our algorithms are tractable to run on modest hardware in real time.

We also investigated the ability to disambiguate multiple simultaneous motion locations. In these experiments, we had multiple people walking in a building and attempt to distinguish between each motion event. We found the base detection results unchanged, and the geometric accuracy and precision of our approach degraded only by a few feet, making our approach robust to a handful of simultaneous events; we leave a study investigating the scalability of simultaneous events as future work.

The remainder of this paper is organized as follows. Section 2 covers some background on signal propagation and related work. Section 3 then describes the detection and localization algorithms. Next, Section 4 documents the accuracy, precision, latency, and geometric accuracy of our approach in two different office environments for single and multiple motion events. Finally, in Section 5 we conclude.

2 Background and Related Work

Our work is founded on two well tested ideas from signal propagation theory. The first concerns how to characterize the multipath environment, and the second how people impact the received signal in such an environment. The key theoretical result we leverage from prior signal propagation work is that the variance in response to human motion should be significantly higher than the variance resulting from noise. A second less used result is that the signal's dominant component is most often a straight line from the transmitter to the receiver [6]. A key result from our work is that simple, intuitive application of these results is sufficient for mobility detection and localization. It remains an open question if, and by how much, more complex models will improve the geometric bounds provided by our approach.

2.1 Background

Modeling the received signal strength (RSS) in indoor environments is challenging task because of many effects on the radio waveform. These include shadowing, *i.e.* blocking a signal; reflection, *i.e.* waves bouncing off an object; diffraction, *i.e.* waves spreading in response to obstacles; and refraction, *i.e.* waves bending as they pass through different mediums. From a system perspective, the modeling challenge is not to provide absolute accuracy, but rather to find a model that balances accuracy, generality, computational complexity and parameter observability for the task at hand.

It has been empirically demonstrated in numerous studies

that indoor signal power distributions are typically Rician. Such distributions model environments where a few dominant paths compose the received signal. This stands in contrast to Rayleigh environments, where a received signal is composed of many uncorrelated components. The important result we draw on from the Rician models is that a change in a few components will have comparatively large effects compared to noise. Thus, the resulting stochastic process will have a high variance, even on short timescales. A second result from previous modeling efforts is that the impact of objects impacts the field in the LOS between a transmitter and receiver [6]. Although more complex geometries, e.g. modeling the multipath components as ellipses rather than rays [6], or modeling ground effects [12] could be used, we found that such complexity did not deliver much benefit over modeling the propagation using straight lines.

2.2 Related Work

One of the first works to articulate the area of passive detection was by Youssef *et al.* [11]. Using terminology from that work, we focus on detection and tracking. However, our work expands on those definitions by making mobility itself a detection event, as opposed to object detection. Most other works assume detection via *static localization*, where the goal is to locate stationary objects. A second class of related works examine *location discrimination*, which determines if an object has changed positions [13]. All three are distinct problems. As the field is still relatively new, there are still few formal problem definitions and accepted operating regimes for those problems. Note that mobility detection and location discrimination are easy problems if we had a static localization system with zero error that operated with zero latency and with an infinitely small time granularity. However, given the current state of the art, mobility detection and location discrimination are useful.

Under certain conditions these problems can be reduced from one to another. For example, mobility detection and location discrimination may or may not be equivalent depending on the latency used when defining the events. For example, suppose a transmitter is moved and placed back in the same location. Under the definition of mobility detection, an event has always occurred, however, a change in location may or may not have occurred depending on the time-granularity used in location discrimination. In the above example, mobility detection and location discrimination become equivalent as the time-granularity approaches zero.

In the realm of motion detection, there has been much work using the variance of the RSS to determine if a wireless device has moved. Wallbaum and Diepolder proposed a simple motion detection scheme based on RSS readings from a Wireless LAN [8]. They used a variance threshold to declare that a device is moving or stationary. With three access points (APs) and a sliding window of size five samples per AP, they got five percent false positives and about ten percent false negatives. Muthukrishnan *et al.* also presented a motion detection algorithm that was based on the spectral analysis of WLAN radio signal strengths by employing Fast Fourier Transform [2, 3]. A two-state classification scheme was used to deduce if a user is moving or still with an average classification accuracy of 94%. Patwari and Kasera proposed a location distinction mechanism that used a physical layer characteristic of the radio channel, called temporal link signature, between a transmitter and a receiver to detect when the transmitter or receiver changed position [5]. Finally, Xing *et al.* explored the use of mobile sensors to address the limitation of wireless sensors networks for target detection [10]. Target detection relies on sensing changes in the energy of signals emitted by targets.

The ability to passively localize humans has been recently approached by several groups. Seifeldin and Youssef, for example, first use an offline training phase, and then apply a matching approach based on Bayes' rule [7]. Their matching approach has its roots in machine learning, and thus unlike our geometric approach which is more an approximation founded on signal propagation. Our results show that for the mobility detection problem the signal maps are sufficiently simple that machine learning approaches are not required.

Radio tomographic imaging seeks to localize object in the environment using the signal properties. Two recent works take this step to the logical conclusion of trying to map a human body using wireless networks [4, 6, 9]. Although our localization could be viewed as an approximation of tomographic imaging, it is important to realize the end goals are much different: we seek only to return a sufficiently constrained area for an event rather than build a model of the object(s) contained in a space, making our modeling and algorithms much simpler. Also, a second key difference arising from the different goals is that the deployment densities and regularity needed for the motion detection and localization problems are an order of magnitude less than for tomographic imaging. For example, Wilson and Patwari's recent work reported densities of 1 transceiver every 15.75 ft^2 [9] compared to our highest density of 1 tag every 250 $ft.^2$.

Another work close to ours in approach is by Zhang *et al.* [12]. They use an intersecting lines approach with wireless sensors to localize moving people. That work also took the approach of deploying a regular grid of sensors at a density much higher than ours. Their approach of using the midpoint of the interesting lines works well if the sensors are located in a regular grid. However, we found the density to be quite variable in actual deploy-

ments, and as a result propose an approach more robust to different deployment densities.

3 Geometric Approach

We now describe our motion detection and localization approach. Our system requires *a priori* knowledge of the locations of the transmitters and receivers being used for passive motion localization. This information can either be provided by an existing localization system, or it can be manually determined by the user/administrator of the system. An infrastructure-based localization system already has knowledge of the location of its passive receivers, and could potentially localize the transmitters being used to provide an estimate (within 2-6 meters) of their locations [1].

The algorithm divides the monitored space into a matrix of overlapping rectangular tiles, **T**, where each tile overlaps half of the geometric space covered by neighboring tiles, except for tiles located at the edges, which only overlap their interior neighbors. By overlapping tiles we avoid aliasing the space along tile boundaries. The intuition is that non-overlapping tiles may incorrectly report the location of motion that occurs along tile boundaries. We find that this approach produces qualitatively better results than using non-overlapping tiles. The other approach to anti-aliasing, shrinking the tile size, resulted in linearly-shaped areas in regions of low density.

Given a set of *n* receivers, denoted $\mathbf{R} = \{r_1, r_2, \ldots, r_n\}$ and a set of *m* transmitters, denoted $\mathbf{S} = \{s_1, s_2, \ldots, s_m\}$ we identify the signal from s_i received by some r_j as $rssi_{ij}$. For each pairwise signal between a transmitter s_i and a receiver r_j, we compute the standard deviation of the signal over a 3-second sliding window, and define this as σ_{ij}. The advantage of using the standard deviation is that it is not dependent upon the actual value of the received RSSI, though long-distance links typically have very low RSSI values (< -80 dBm), which can be influenced by more environmental effects than short-range links. To overcome this effect, we calculate a circle with its center at the center of a tile, and some radius T_{Rad}. For our experiments a radius of 90 ft. was used, which was experimentally determined to reduce fluctuations caused by very long links while not discarding large amounts of data significant to localizing motion.

For each tile t_k we discard any σ_{ij} that meets any of the following criteria:

1. the value of σ_{ij} is below a threshold value, σ_T;
2. either endpoint of σ_{ij} (s_i or r_j) is farther than the radius, T_{Rad}, from the center of t_k; or
3. the line segment connected s_i and r_j does not pass through the interior of t_k.

For all σ_{ij} remaining, the tile's score is computed using the following formula:

$$\sum \left[\frac{(\sigma_{ij} - \sigma_T)}{d_{ij}^{(d_p)}} \right] \quad (1)$$

where d_{ij} is the length of the line segment connecting s_i and r_j.

Intuitively, tiles receive a higher score when the standard deviation of the intersecting lines is higher than a base noise level σ_T. This base score is then discounted by the length of the line. We used an exponential discounting factor because signal power is logarithmic with distance, thus long lines receive much less weight than short ones. One can view this as a very gross approximation of more sophisticated models, e.g., Cassini ovals with exponential path loss [6].

Table 1 shows the single set of "best" values for the constants in our approach. We determined this set by performing a brute-force search through the parameter space maximizing both precision and recall from a single trace collected in the CoRE hall building. We then collected additional traces at CoRE hall varying the tag/receiver densities, and also at WINLAB. We found this set of constants delivered good performance for these environments. Given the good performance over a range of environments, our approach is unlikely to need a lot of tuning as the environment changes. We leave investigation of the sensitivity of our algorithm with respect to these constants, and as a function of different environments, as future work.

Parameter	Value
Tile matrix	19 x 7
T_{Rad}	90ft.
σ_T	1.2
d_p	1.1
$Ratio_N$	0.7
$Ratio_P$	0.5
D_T	0.5 (0.2 for 2 traces)

Table 1: Parameter values used to detect and localize motion in the two experimental spaces.

The scored tiles are then sorted in non-decreasing order according to score. If the score of the highest-scored tile ("peak tile") is greater than the detection threshold D_T, a depth-first recursive cutting algorithm is applied to "remove" neighboring tiles that have lower scores than the peak tile. In our approach, a "cut" tile has its score set to 0. Given a peak tile, T_{Peak}, the current tile being inspected for cutting, T_0 (initially, this will also be the peak

Space	Dimensions	Tile Grid	Tile Size
CoRE hall	220ft. x 88 ft. ($19,000ft.^2$)	19 x 7 (133)	22ft. x 22 ft. ($484ft.^2$)
WINLAB	164ft. x 74 ft. ($12,000ft.^2$)	19 x 7 (133)	16ft. x 18 ft. ($288ft.^2$)

Table 2: Dimensions of the two experimental spaces. CoRE hall is a large 7 story building, of which the third floor houses the computer science department. The WINLAB building a section of a large single story building housing the WINLAB research center.

tile), and some neighboring tile, T_{Nbr}, the neighbor tile is cut under the following conditions:

1. T_{Nbr}'s score is less than D_T;
2. $\frac{score(T_0)}{score(T_{Nbr})}$ is less than the "neighbor ratio", defined $Ratio_N$;
3. $\frac{score(T_{Nbr})}{score(T_{Peak})}$ is less than the "peak ratio", defined $Ratio_P$.

We continue cutting neighboring tiles in a depth-first approach until a neighboring tile's score is greater than the current tile's score. Although this is an *ad-hoc* approach to finding local maxima in the matrix of tiles, we found it works well for our purposes. Future work will explore the possibility of using more traditional search methods for determining local maxima.

Any tile checked and not cut while performing the depth-first traversal of neighboring tiles are added to the peak tile's *tile group* (TG). The results of motion localization are returned as one or more tile groups, depending on the number of independent motion events detected by the system. The scoring, sorting, and cutting of tiles generates a single tile group indicating the likely location of motion. All σ_{ij} that contributed to the scores of the tiles in this tile group are removed from the global set, and the process of scoring and cutting tiles is repeated to generate a new tile group. Repeating this process enables our system to detect multiple areas of motion if they are sufficiently distant from each other or caused by disjoint sets of lines and tiles. Once this repeated process no longer returns a new tile group, the system reports the one or more tile groups generated. Figure 1 provides an illustration of the tag/receiver lines, a single tile group reported by our motion localization system, and two ground truth motion events: one that is "covered", and one that is not.

As a result of our geometric approach, the system is somewhat sensitive to the size of the tiles with respect to the size of the area, the number of tags and receivers, and the density of their placement. Very large tile sizes will return large areas of motion, while very small tile sizes (1-2 feet) will essentially be "pixelized" representations of the σ_{ij} line segments used to score the tiles. In addition, the system is sensitive to the accuracy of the locations of the transmitters and receivers with respect to the tile size. With large tile size ($400ft.^2$), a location error of 1-3 feet will be unlikely to significantly affect the results of motion localization, while small tile sizes ($25ft.^2$) are less robust to measurement errors. During the brute-force parameter space search we also investigated different tile sizes, finding the 19x7 tiling listed in Table 1 to be a balance between higher-precision, smaller tile sizes and computational limits for real-time application of the algorithm.

4 Experimental Method and Results

In this section, we first describe our experimental approach. We then describe the metrics we use, but make the formal definitions in later sections. For each metric, we first describe the results using a single person generating mobility events. We then describe the impact of latency and multiple people generating events.

4.1 Approach

We performed experiments in two different spaces. The first space is one floor of CoRE hall, which is a large office building (220ft. x 88ft), of which the third floor is dedicated to the Computer Science Department. We refer to this space as B_1. The second space is a section of a research lab (164ft. x 74ft.), and this houses the Rutgers WINLAB, and we call this B_2 (Table 2). We deployed receivers and tags throughout both spaces, attempting to provide effective coverage of the areas where traces would be collected. Because CoRE hall contains private offices around the perimeter of the building, we were unable to place receivers around the perimeter except in two conference rooms. WINLAB was more accessible, and we were able to place some receivers around the perimeter of the experimental space, but were still somewhat constrained by the availability of networking and power connections. Figure 2 contains diagrams for the 3 deployments in CoRE hall and the deployment of WINLAB.

Within each experimental space, we configured a localization system along with a set of active RFID tags [1] with receivers at known locations. We then deployed a set

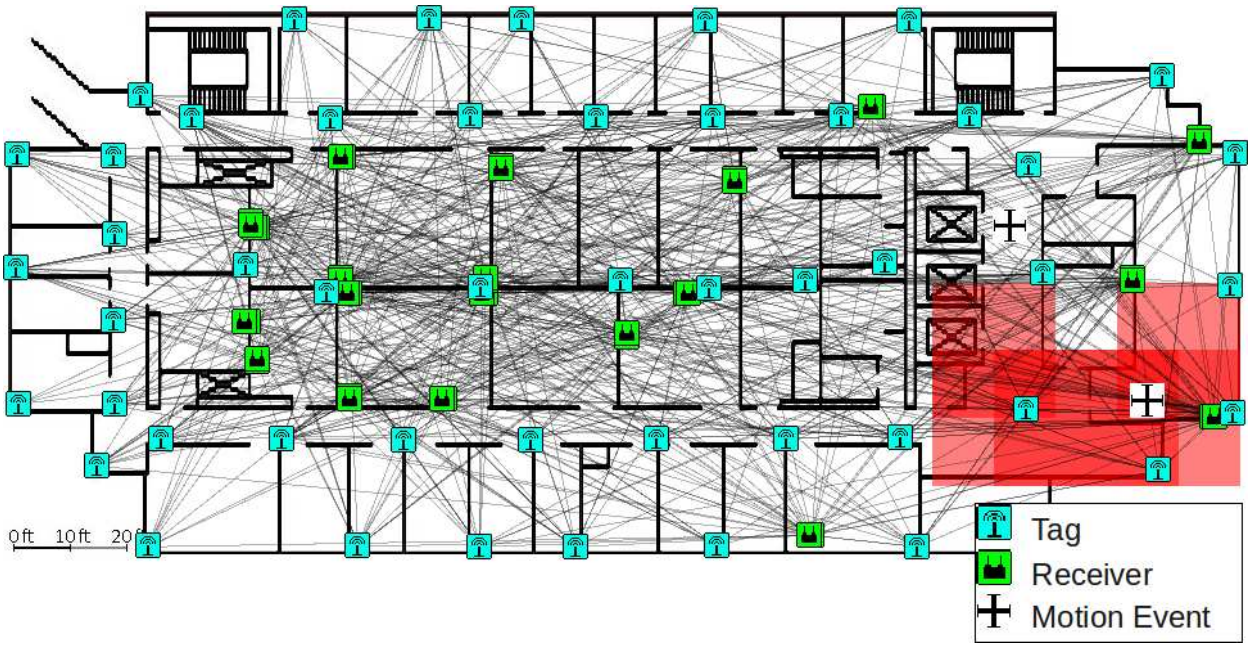


Figure 1: Image of building 2 with densest tag deployment. Tags and receivers are identified by small square icons, the signal lines are drawn, and several tiles can be seen to cover a ground truth crosshair in the lower-right quadrant.

of tags throughout the space and measured their locations. The readers are constantly listening for beacons from the tags, and the tags transmit beacons at 1-second intervals.

A motion localization solver was connected to the localization server and the server periodically sends the standard deviation of the last 3 seconds of RSSI data available for each transmitter/receiver pair. For our experiments, the period was set a 1 second, though this could vary depending on system constraints such as network latency, bandwidth, or server/solver hardware performance.

For each experiment, we asked the participants to walk along a designated path through the space with a paper notebook and a stopwatch. At a number of predefined landmark locations, the participants would record the current reading of the stopwatch. After the signal data was collected for an experiment, the locations of the participants were entered manually into a data file for each time point that signal data was sent to the motion localization solver. For time points where no fixed location was known (*i.e.* the participant was not at a landmark location), the positions were manually interpolated and entered. The longest time interval between landmark locations was less than 20 seconds, which would require the interpolation of 18 points along a straight path.

For each experimental space, we also collected data during "quiet" periods where no motion was taking place (Table 3). CoRE hall has 3 elevators present, identifiable in the diagrams as crossed boxes on the right side (Figure 2), and since it was not possible to disable the operation of the elevators during the experiments, this is a potential source of experimental error and the impact of unregistered motion is not addressed in this paper. Any unknown motion occurring in either space would be evaluated as false positives.

Trace Description	FP	TE	FP %
B_1, Ext. Only, $D_T = 0.2$	12	113	10.6%
B_2, $D_T = 0.5$	1	144	0.7%
B_2, $D_T = 0.5$	12	852	1.4%

Table 3: False positive rates for traces collected in the two experimental spaces. Unregistered motion or the necessarily low D_T (resulting from the sparse tag deployment) for CoRE hall may be the cause of increased false positive rates.

4.2 Evaluation Metrics

We characterize performance along several different dimensions. The most important axes of characterization are detection accuracy (recall, precision), spatial accuracy (coverage rate/accuracy), and temporal accuracy (latency). The first set of metrics focus on detection and view mobility detection in a binary manner: either the system detects a mobility event or does not (Section 4.3). The second set of metrics focuses on the spatial accuracy: the first spatial metric describes how well the tiles cover the

Metric	Formula
Recall	$\frac{TP}{TP+FN}$
Precision	$\frac{TP}{TP+FP}$
Coverage Rate (CR)	$\frac{GT_C}{GT_T}$
Coverage Accuracy (CA)	$\frac{T_C}{T_T}$

Table 4: Definition of metrics used to evaluate the performance of the motion localization algorithm.

mobility events (Section 4.4; the other characterizes the sizes of the returned area for both covered and uncovered mobility events (Section 4.5). We then examine the impact of latency on the above metrics (Section 4.6). Finally, we describe how these metrics hold up in simultaneous multi-event situations.

For evaluating the performance of our motion localization system, we use a sliding time window of 2-5 seconds for ground truth locations. This is because the server and network performance can delay processing or transmission of data, and also because the standard deviation is computed over a 3-second window and may not change significantly until 1 second after the change in RSSI indicates motion between a sender/receiver pair.

We further evaluate the performance of the motion localization system by investigating the following quantities: the geometric error of both covered ground truths and and non-covered ground truths, and the time latency of covered ground truths.

4.3 Motion Detection

We define the following metrics in the context of motion detection: *true positive*, *false positive*, *true negative*, and *false negative*. A *true positive* (TP) is declared at some point in time when there exists at least one ground truth and at least one tile group returned by the motion localization system. This is a binary decision and enables the results of our work to be compared with previous work in motion detection systems. A *false positive* (FP) occurs when the motion localization system returns one or more tile groups when there are no ground truths present at some point in time. A *false negative* (FN) is defined as the motion localization system returning no tile groups when there is at least one ground truth present at some point in time. A *true negative* (TN) is reported when the motion location system reports no tile groups and there are no motion ground truths present at some point in time.

We define recall and precision in the standard way, reproduced in Table 4. Further discussion and explanation of Coverage Rate (CR) and Coverage Accuracy (CA) is contained in Section 4.4.

The results for each experiment are provided in Table 5 for a single motion event and Table 6 for multiple motion events. As a motion detection system, our approach produces relatively good results with recall above 90% for every trace except 1, and precision above 90% for all experiments. False positives typically centered around areas with a very high density of lines intersecting tiles, and reducing the false positive rate will be investigated in future work. The reason for the low recall of $B_1(3)$ is due to a "blind spot" caused by the tag arrangement (Table 5). This results in a relatively large number of false negatives for the experiment, pushing the recall below 0.85. The same motion path was followed in $B_1(6)$, and due to denser tag placement in the affected area, the recall and coverage rate metrics were improved, though small blind spots were still observed.

For most of the experiments listed, the detection threshold was set to 0.5, though for $B_1(8)$, the detection threshold was set to 0.2 due to the lower density and increased distance of tags from the actors walking through the space.

4.4 Tile Coverage

Motion localization is a more difficult task than detection, and we introduce two metrics to evaluate our system's performance in this respect: *coverage* is defined as when a tile group geometrically encapsulates a ground truth position, and *non coverage* is defined as when a tile group fails to encapsulate a ground truth position. The term encapsulate means that a ground truth is contained within the boundaries of one or more tiles within a tile group. We further define the metric *coverage rate* (CR) as the number of ground truth motion events that are covered by a tile group (GT_C) divided by the total number of ground truths (GT_T). Finally, we define *coverage accuracy* (CA) as the number of tile groups covering a ground truth motion event (T_C) divided by the total number of tile groups returned by the system (T_T).

The motion localization portion of our algorithm is evaluated in the coverage rate (CR) and coverage accuracy (CA) columns in Tables 5 and 6. For both metrics, we report both the instantaneous values, the values obtained when allowing a 2-second and 5-second sliding window for ground truth motion events. The 5-second sliding window, as described in Section 4.1, is used because of delays inherent in both the data reporting and the calculation of the standard deviation.

The coverage rate (CR) column indicates the number of ground truth motion events covered by some tile set when using a window size of 0 seconds. The coverage rate metric provides a useful metric in determining how well the localization portion of the algorithm performs. The coverage rate for a window size of 2 and 5 seconds is also provided, and the significant improvements indicate how

Trace	Tag Placement	Recall	Precision	CR	CR (2s)	CR (5s)	CA	CA (5s)	Lat. (90%)	Lat. (100%)
$B_1(1)$	Corridors	0.935	0.980	0.867	0.929	0.959	0.864	0.961	1.1s	3.4s
$B_1(2)$	Corridors	0.923	0.980	0.843	0.921	1.000	0.808	0.960	1.2s	3.4s
$B_1(3)$	Corridors	0.849	0.964	0.784	0.829	0.865	0.830	0.938	$<$ 1s	4.5s
$B_1(4)$	Corr. & Ext.	0.949	1.000	0.899	0.949	1.000	0.890	0.910	1.2s	4.5s
$B_1(5)$	Corr. & Ext.	0.958	1.000	0.926	0.958	0.968	0.817	0.844	$<$ 1s	2.2s
$B_1(6)$	Corr. & Ext.	0.905	1.000	0.760	0.913	0.971	0.800	0.950	1.2s	3.4s
$B_1(7)$	Corr. & Ext.	0.918	1.000	0.855	0.936	1.000	0.862	0.923	1.2s	4.5s
$B_2(1)$		0.984	0.984	0.930	0.982	1.000	0.757	0.824	$<$ 1s	2.1s
$B_2(2)$		0.955	1.000	0.917	0.983	1.000	0.781	0.849	$<$ 1s	2.1s
$B_2(3)$		0.983	1.000	0.857	0.929	1.000	0.762	0.889	1.1s	2.3s
$B_2(4)$		0.984	1.000	0.898	0.966	1.000	0.871	0.919	$<$ 1s	2.1s

Table 5: Detection rates for single motion events in the two experimental spaces. Receiver and tag placement for Buildings 1 and 2 are shown in Figure 2.

Trace	Tag Placement	Actors	Recall	Precision	CR	CR (2s)	CR (5s)	CA	CA (5s)	Lat. (90%)	Lat. (100%)
$B_1(8)$ [$D_T = 0.2$]	Exterior	2	1.000	0.938	0.797	0.879	0.970	0.872	0.936	1.2s	4.9s
$B_1(9)$	Corr. & Ext.	3	0.991	1.000	0.807	0.905	0.982	0.757	0.831	1.3s	4.9s
$B_1(10)$	Corr. & Ext.	3	0.984	1.000	0.725	0.815	0.911	0.689	0.762	2.4s	4.8s
$B_1(11)$	Corr. & Ext.	3	0.983	1.000	0.834	0.915	0.964	0.770	0.808	1.3s	4.9s
$B_1(12)$	Corr. & Ext.	3	0.982	1.000	0.748	0.834	0.945	0.710	0.771	2.4s	4.8s
$B_2(5)$		2	1.000	1.000	0.860	0.926	0.971	0.901	0.959	$<$ 1s	4.2s
$B_2(6)$		3	0.984	1.000	0.806	0.893	0.961	0.835	0.924	1.1s	4.0s
$B_2(7)$		3	0.987	1.000	0.775	0.887	0.979	0.890	0.990	1.2s	4.2s

Table 6: Detection rates for multiple motion events in the two experimental spaces.

latency in the system affects the performance. Using a 5-second window size, the coverage rate for all experiments (except $B_1(3)$) is consistently above 90% and demonstrates that the system can effectively capture the locations of motion, though not without a modest time latency.

The coverage accuracy (CA) column indicates what percentage of tile groups reported by the algorithm cover a ground truth motion event when using a window size of 0 seconds. The coverage accuracy metric provides a way of determining how likely a given tile group is to contain a ground truth motion event. With a window size of 5 seconds, the coverage accuracy shows modest improvements, though not as significant as the coverage rate. With the 5-second window, coverage accuracy values are above 76%, with a median of 91.9%. For single motion, the median coverage accuracy is also 91.9%, while for multiple motion events, the median value drops to 87.8%.

The latency (Lat.) columns indicate that 90% of the ground truth motion events are covered within 1.2 seconds, and 100% within 5 seconds, though this is the maximum latency evaluated for our system. For the single motion experiments, the 90^{th} percentile is always less than or equal to 1.2s, while for two of the multiple motion experiments this increases to 2.4 seconds. Contrasting the different experimental spaces, the smaller, higher-density WINLAB overall has lower latencies than CoRE hall, implying that higher tag densities can decrease time delays in detection. It should also be noted that WINLAB allowed better placement of receivers, resulting in a more even and thorough coverage of the experimental space, translating into lower coverage latencies.

4.5 Distance Accuracy

In this section we characterize how close a set of tiles are to the ground truth. Our motion localization algorithm typically returns somewhat regular groups of tiles, meaning that the tile group itself forms a semi-rectangular shape with a few outliers or "tails". Figures 3 and 4 show cumulative distribution function (CDF) curves for the localization error in CoRE hall and WINLAB, respectively. Each graph contains 6 plots divided into 2 categories: "coverage", tile groups that cover one or more ground truth motion events, and tile groups that do not cover any ground truths (Section 4.4). When a tile group covers a ground truth, the distances between the ground truth and the centers of all tiles in the tile group are calculated. The minimum, median, and maximum distances from the tile group are then plotted in the graph respectively as "Cover [Min]", "Cover [Med]", and "Cover [Max]". For tile groups that do not cover any ground truth events, the distance between the center of the highest-scoring tile in the

group and all ground truths are calculated and the nearest ground truth is selected. The distances for all tiles in the group are computed in the same way as for covering tile groups, and the results are shown in the graphs.

By looking at a horizontal slice at the 50^{th} percentile point in the CDF graphs, an approximation of the width/height of the tile groups can be inferred. Using Figure 4b as an example, making a horizontal cut across the 50^{th} percentile point shows that for both covering and non-covering tile groups, the minimum distance is around 4 ft. and the maximum around 20 ft. This would indicate that the median tile group for this trace is approximately 16 on a side. Given that the tiles are approximately 16 ft. x 18 ft. and overlap each other, the median tile group returned for this trace is likely to be composed of approximately 3x3 tiles.

It can also be seen by comparing Figures 3 and 4 and looking at the error scales on the horizontal axes that WINLAB's smaller space generally performs better in terms of distance accuracy. This is likely a combination of factors including a more open environment (less obstructions and few walls), a denser deployment, and tag deployment that provides more even coverage than that used in CoRE hall.

4.6 Impact of Latency

As seen in Tables 5 and 6, latency is a significant factor. The instantaneous localization results (CR and CA columns) demonstrate that for many of the experiments, the localization of motion events is significantly improved by allowing for a time window of 5 seconds. Coverage rate in particular increases as the window is increased, and coverage accuracy to a lesser extent. The effects of the time window on coverage rate become apparent by realizing that a sliding window of 5 seconds essentially provides the algorithm with 5 "opportunities" to cover the actual mobility event.

With only a 2-second window, the coverage rate increases by 50% or more of the difference between the 5-second window and the instantaneous results. These results agree with the two latency columns (Lat. 90% and Lat. 100%), which show that for single motion experiments, and multi-motion experiments to a lesser extent, the majority of the motion events are covered by tile groups returned by the algorithm with 1.2 seconds of the actual event (1.3 seconds for multiple motion). These experimental results agree with the analysis in section 4.1, which argues that latency is a factor both of physical limitations (network latency, data processing) and of mathematical limits imposed by using standard deviation of signal data as a basis for motion localization.

4.7 Multi-User Detection and Localization

Looking at Tables 5 and 6, we observe that when multiple actors are simultaneously moving through a space, the overall amount of interference generated is much greater than for a single actor, so motion detection becomes more reliable. This can be seen by comparing the recall and precision rates for single motion and multiple motion. The outlier in the multi-user traces is $B_1(8)$ where the precision is the lowest out of all the experiments, an artifact of the low density of tags and larger distance between the moving actors and the tags. For this experiment, tags were placed on the exterior of the building, inside offices with closed doors, approximately 15-20 ft. from the actual motion. As a result, the detection threshold was set lower, which increased the false positive rate and lowered precision.

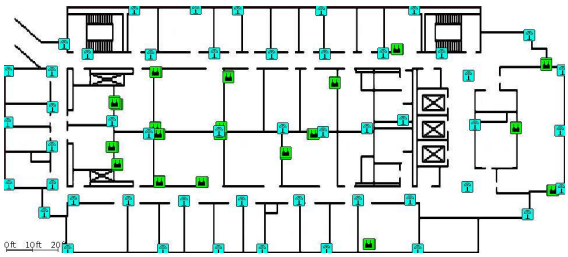
Multiple simultaneous motion also makes it more difficult for the system to determine the likely location of the motion, since the algorithm must separate multiple events from a global set of data. This is demonstrated by the much lower coverage rates and accuracy for multiple motion events contrasted with single motion. The median instantaneous coverage rate for single motion is 0.867 while for multiple motion, it is .802. With a 5-second window, the median coverage rate for single motion is 1.0, and for multiple motion is 0.967.

5 Conclusions

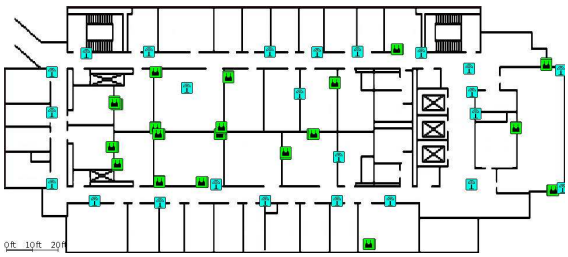
In this work we demonstrated very high accuracy and precision for passive mobility detection in indoor environments. We showed the feasibility of localizing the motion to within 10-15 ft using a live system in two different buildings. We also determined that it is possible to obtain such performance on the timescales of a few seconds, with modest computing hardware requirements, and with active RFID tag and reader densities on the order of 1/500 ft.^{2}. We also showed that this level of performance can be achieved by utilizing only the standard deviation of the received signal strength and the geometry of the receivers and tags. Neither complex propagation models nor any training data are required.

Our work also leaves several open problems. The most immediate is improving the geometry of the mobility localization. Using more sophisticated propagation models it may be possible to constrain the sizes of the areas more than we do now. It may also be possible to reduce the worst-case error, which is still substantial. A second approach might be to use more temporal reasoning, *i.e.* tracking, in order to constrain the possible areas or reduce the maximum error.

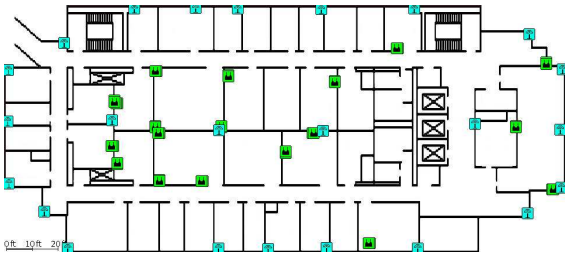
Another implication of our work is that it points to a



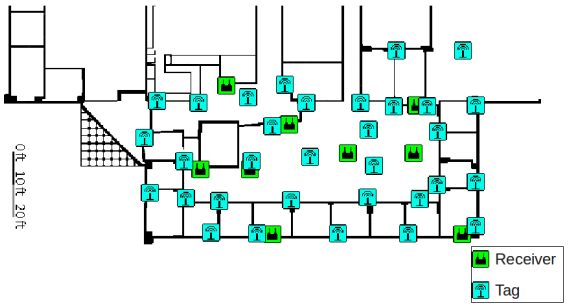
CoRE hall with the tags on corridor walls and exterior offices. Used in $B_1(4-7, 9-12)$.



CoRE hall with tags only on corridor walls. Used in $B_1(1-3)$.

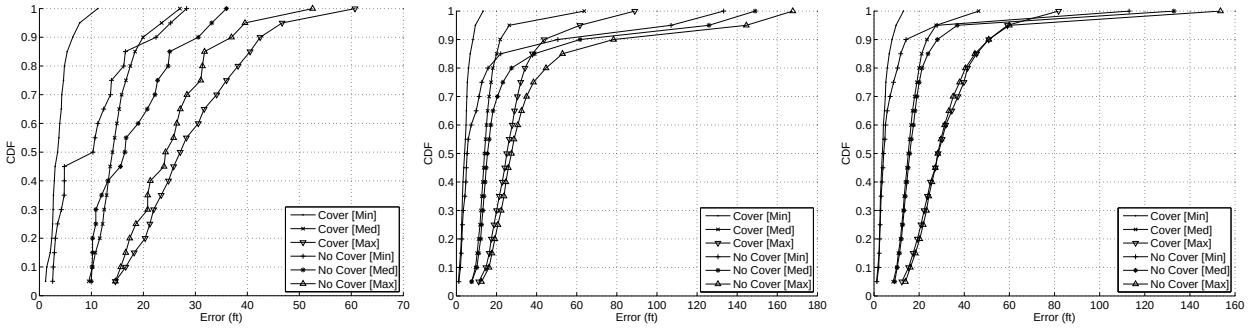


CoRE hall with tags only on exterior offices. Used in $B_1(8)$.



WINLAB with tag locations. Used in $B_2(1-7)$.

Figure 2: Tag and receiver deployments for the two experimental spaces. The bottom-right space is WINLAB, the three others are CoRE hall.



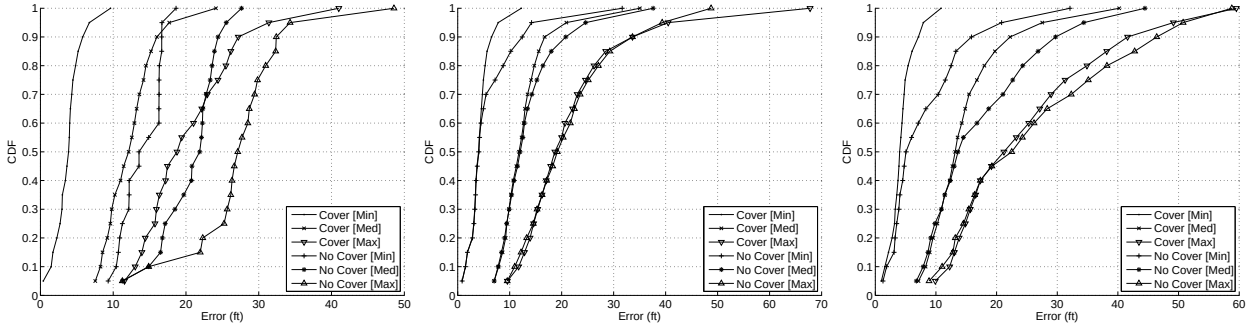
(a) 1 Person (b) 2 People (c) 3 People

Figure 3: Geometric distance error for 3 traces in CoRE hall (large office space). 3a ($B_1(4)$) and 3c ($B_1(11)$) have tags on the exterior office walls and also in the corridors. 3b ($B_1(8)$) has tags only on the exterior office walls.

broader use of wireless networks for mobility estimation. As technology trends enable the inclusion of wireless networks into every communication device, the resulting density and communication volume increases will make the background radio spectrum sufficiently saturated to observe the fluctuations resulting from object motion. Ideally, the wireless infrastructure could provide high precision passive motion detection and medium scale mobility localization with little additional cost over the base communication system. Our work makes one small step towards realizing a general purpose layer that tracks passive motion, however, much more work on higher layers of software are needed to realize this vision in practice.

References

[1] Y. Chen, E. Elnahraway, J.-A. Francisco, K. Kleisouris, X. Li, H. Xue, and R. P. Martin. Grail: General Real Time Adaptable Indoor Localization. In *Proceedings of the 4th ACM Con-*



(a) 1 Person (b) 2 People (c) 3 People

Figure 4: Geometric distance error for 3 traces in WINLAB (small office). From left to right, $B_2(4)$, $B_2(5)$, and $B_2(7)$.

ference on Embedded Networked Sensor Systems (SENSYS), Demo Abstract, Nov. 2006.

[2] K. Muthukrishnan, M. Lijding, N. Meratnia, and P. Havinga. Sensing Motion Using Spectral and Spatial Analysis of WLAN RSSI. In *Proceedings of the Second European Conference on Smart Sensing and Context (EuroSSC 2007)*, pages 62–76, October 2007.

[3] K. Muthukrishnan, B.-J. van der Zwaag, and P. J. M. Havinga. Inferring motion and location using wlan rssi. In *The Second International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments (MELT)*, pages 163–182, Orlando, FL, USA, Sept. 2009.

[4] N. Patwari and P. Agrawal. Effects of correlated shadowing: Connectivity, localization, and rf tomography. In *Proceedings of the International Conference on Information Processing in Sensor Networks (IPSN)*, pages 82–93, St. Louis, Missouri, USA, Apr. 2008.

[5] N. Patwari and S. K. Kasera. Robust Location Distinction using Temporal Link Signatures. In *Proceedings of the ACM International Conference on Mobile Computing Networking (Mobicom 2007)*, pages 111–122, September 2007.

[6] N. Patwari and J. Wilson. People-sensing spatial characteristics of rf sensor networks. *CoRR*, abs/0911.1972, 2009.

[7] M. Seifeldin and M. Youssef. Nuzzer: A large-scale device-free passive localization system for wireless environments. *CoRR*, abs/0908.0893, 2009.

[8] M. Wallbaum and S. Diepolder. A motion detection scheme for wireless lan stations. In *Proceedings of the 3rd International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*, London, UK, Oct. 2006.

[9] J. Wilson and N. Patwari. Radio tomographic imaging with wireless networks. *IEEE Transactions on Mobile Computing*, 2010. to appear.

[10] G. Xing, J. Wang, K. Shen, Q. Huang, X. Jia, and H.-C. So. Mobility-assisted spatiotemporal detection in wireless sensor networks. In *28th IEEE International Conference on Distributed Computing Systems (ICDCS)*, pages 103–110, Beijing, China, June 2008.

[11] M. Youssef, M. Mah, and A. K. Agrawala. Challenges: device-free passive localization for wireless environments. In *Proceedings of the 13th Annual International Conference on Mobile Computing and Networking (MOBICOM)*, pages 222–229, Montréal,Québec, Canada, Sept. 2007.

[12] D. Zhang, J. Ma, Q. Chen, and L. M. Ni. An rf-based system for tracking transceiver-free objects. In *Proceedings of the Fifth IEEE International Conference on Pervasive Computing and Communications (PERCOM)*, White Plains, NY, USA, Mar. 2007.

[13] J. Zhang, M. H. Firooz, N. Patwari, and S. K. Kasera. Advancing wireless link signatures for location distinction. In *Proceedings of the 14th Annual International Conference on Mobile Computing and Networking (MOBICOM)*, pages 26–37, San Francisco, CA ,USA, Sept. 2008.