

Distinguishing Locations Across Perimeters Using Wireless Link Measurements

Junxing Zhang

Email: junxing@cs.utah.edu

Sneha K. Kasera

Email: kasera@cs.utah.edu

Neal Patwari

Email: npatwari@ece.utah.edu

Piyush Rai

Email: piyush@cs.utah.edu

Abstract—Perimeter distinction in a wireless network is the ability to distinguish locations belonging to different perimeters. It is complementary to existing localization techniques. A drawback of the localization method is that when a transmitter is at the edge of an area, an algorithm with isotropic error will estimate its location in the wrong area at least half of the time. In contrast, perimeter distinction classifies the location as being in one area or the adjacent regardless of the transmitter position within the area. In this paper, we use the naturally different wireless fading conditions to accurately distinguish locations across perimeters. We examine the use of two types of wireless measurements: received signal strength (RSS) and wireless link signature (WLS), and propose multiple methods to retain good distinction rates even when the receiver faces power manipulation by malicious transmitters. Using extensive measurements of indoor and outdoor perimeters, we find that WLS outperforms RSS in various fading conditions. Even without using signal power WLS can achieve accurate perimeter distinction up to 80%. When we train our perimeter distinction method with multiple measurements within the same perimeter, we show that we are able to improve the accuracy of perimeter distinction, up to 98%.

I. INTRODUCTION

Perimeter distinction in a wireless network is the ability to distinguish locations that are separated by a perimeter or more generally, belong to different perimeters. Unlike location estimation, perimeter distinction does not attempt to determine the exact location of a transmitter, instead, it only attempts to find out to which perimeter a transmitter belongs. Recently, researchers have developed methods for location distinction to tell the locations of two wireless transmitters apart [1], [2]. However unlike location distinction, perimeter distinction distinguishes different locations across perimeters instead of any location spots. Thus, perimeter distinction falls between the two extremes of location distinction and location estimation as indicated in Fig. 1.

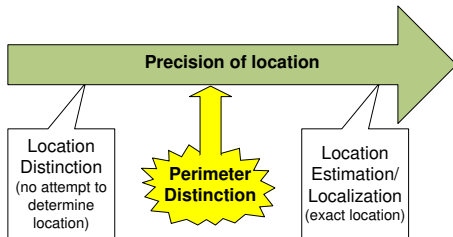


Fig. 1. Spectrum of location distinction, perimeter distinction, and localization in terms of precision of location.

Perimeter distinction has many applications because perimeters are widely used to separate safe zones from unsafe zones, covered areas from uncovered areas, and guarded regions from unguarded region. In secure facilities or buildings, it is important to distinguish between those inside and those outside the building. In supervised living facilities, people who are mentally retarded, or physically handicapped are often required to live in a restricted area. Many parents want to ensure their kids are either inside the house or outside the house, or in one room instead of another. All of these applications, will significantly benefit from the accurately distinguishing locations across perimeters.

Perimeter distinction is complementary to existing RSS-based localization techniques. Regardless of the accuracy of a localization method, when a transmitter is at the edge of an area, an algorithm with isotropic error will estimate its location in the wrong area at least half of the time. Perimeter distinction classifies the location as being in one area or the adjacent regardless of the transmitter position within the area. Furthermore, localization errors in indoor RSS-based localization techniques are comparable to typical room diameters. For example, if the median localization error is 10 feet (the lower bound reported in the comparative study [3]), then a transmitter in the center of a room of diameter 20 feet will be located outside of the room half of the time. Note that this is the best case – as the transmitter position moves away from the center, the performance degrades. From a user’s perspective, a system that locates an object to the incorrect room more than half of the time will be considered an unreliable system.

In this paper, we use the naturally different wireless fading conditions in urban indoor and outdoor environments to accurately distinguish locations across perimeters. Since PHY-layer wireless measurements are products of the fading condition, these can be used for perimeter distinction. We examine the use of two types of wireless measurements for perimeter distinction - one based on the received signal strength (RSS), and another based on the channel impulse response, called wireless link signature (WLS) in the current literature [2]. We also adopt four classification algorithms from statistical learning to classify perimeters.

We consider two scenarios for perimeter distinction. In the first scenario, the wireless receiver discriminates wireless measurements from two sides of a perimeter while considering signal power. In a more complex scenario, the receiver distinguishes measurements from multiple perimeters such

that the distinction is immune to power manipulation in the wireless channel. We compare performance of the two measurement types under the simple scenario and propose multiple methods to retain good distinction rates even when the receiver faces power manipulation by malicious transmitters. We build a methodology that uses signal power as well as multipath characteristics to classify perimeters. Using extensive measurements of indoor and outdoor perimeters, we find that WLS outperforms RSS in various fading conditions. Even without using signal power WLS can achieve accurate perimeter distinction up to 80%. When we train our perimeter distinction method with multiple measurements within the same perimeter, we show that we are able to improve the accuracy of perimeter distinction, up to 98%.

The rest of this paper is organized as follows. In the next section, we describe the application models of perimeter distinction. Section III outlines a framework of perimeter distinction, and also introduces classification algorithms and measurement types used by the framework. We investigate important issues in applying wireless measurements to perimeter distinction in Section IV. Section V evaluates the proposed methods with extensive measurements. The related work are discussed in Section VI. Finally, in Section VII, we conclude the work and indicate directions for future research.

II. APPLICATION MODELS

In this section we start out with a basic model and then extend it to incorporate more complex situations.

A. Basic Model

The basic model for perimeter distinction considers wireless measurements from two sides of a perimeter. In an example, depicted in Fig. 2, two rooms are divided by a shared wall. We consider the scenario where the wireless transmitter (marked as TX in the figure) moves around between two sides of the perimeter while the receiver (denoted as RX) where wireless characteristics are measured, stays at a fixed location at one side and wants to know which side of the perimeter the transmitter currently resides. We also assume in the basic model that the transmitter is incapable of manipulating the signal power for the purpose of fooling the receiver.

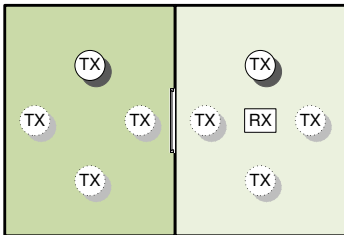


Fig. 2. Basic Model of Perimeter Distinction

B. Advanced Model

In the advanced model we consider multiple perimeters. The example in Fig. 3 shows three rooms divided by two

walls. The leftmost room is separated from the rightmost room by two perimeters whereas other pairs of rooms are segregated by one perimeter. We consider the scenario where the wireless transmitter moves among multiple perimeters while the receiver remains at a fixed location in one perimeter and wonders which perimeter the transmitter stays at the moment. In the advanced model the transmitter can manipulate transmitting power at will in order to trick the receiver.

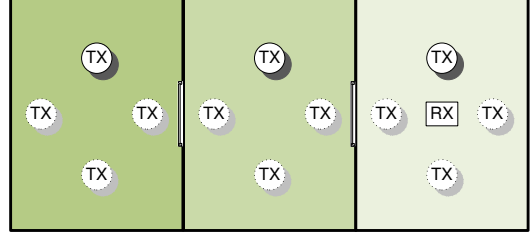


Fig. 3. Advanced Model of Perimeter Distinction

III. FRAMEWORK, CLASSIFIERS, AND TYPES OF MEASUREMENTS

This section first outlines a general framework for perimeter distinction. This framework can flexibly incorporate different classifiers that are described next. The classifiers, or classification algorithms, are adopted from statistical learning [4]. They can categorize unknown measurements into one of the measurement groups corresponding to a specific perimeter. At the end of this section, we introduce the measurement types that are used in our framework to assess fading conditions in perimeters.

A. Algorithmic Framework

We treat the perimeter distinction problem as a supervised learning [4] process for classification. Here, we outline a step-by-step algorithmic framework to allow different classification algorithms to be plugged in for the purpose of perimeter distinction.

- 1) To distinguish perimeters, the receiver first assigns a label $g \in \{1, 2, \dots, M\}$ to each perimeter, where M is the number of perimeters. Next, the receiver must be placed at a suitable location where it can acquire wireless measurements from different perimeters and later perform perimeter distinction tests. Then, the receiver must decide the best feature set of the input measurement vector $\mathbf{h} \in \mathbf{R}^d$ where \mathbf{R} is the set of real numbers and d is the number of features. We discuss the choice of the receiver location and that of the feature set later in Section IV.
- 2) For each perimeter g , the receiver builds a training set of N_g wireless measurements when the transmitter travels inside that perimeter. The receiver stores all measurements and perimeter labels as a training history $\mathcal{H} = \{\mathbf{h}^{(i)}, g^{(i)}\}_{i=1}^N$, where $N = \sum_{g=1}^M N_g$.
- 3) When a new measurement $\mathbf{h}^{(N+1)}$ from an unknown perimeter is obtained, the receiver calls the classification

function with the input measurement and the training history to find out the perimeter label of the new measurement:

$$g^{(N+1)} = \text{classifier}(\mathbf{h}^{(N+1)}, \mathcal{H})$$

The function $\text{classifier}()$ can be any applicable classification functions. We will introduce four such functions in Section III-B.

- 4) The receiver reports the perimeter label output by the classifier. To cope with temporal fading changes in each perimeter and to keep memory usage within a limit, the receiver discards the oldest measurement of the perimeter $g^{(N+1)}$ and adds the latest mapping pair $\{\mathbf{h}^{(N+1)}, g^{(N+1)}\}$ to the history \mathcal{H} . The algorithm returns to step 3 to continue to perform perimeter distinction for new locations.

Although the supervised learning process generally assumes that the training and test data are drawn under similar circumstances, i.e. piecewise stationarity, our perimeter distinction does not rely on this assumption. The non-stationarity can be addressed using simple domain adaptation techniques such as [5], which can be used as a wrapper method around standard classification algorithms in this paper [6]. The history \mathcal{H} can be polluted by the wrong distinction results through updating. Since distinction errors are sporadic while temporal fading changes are consistent [2], continuously updating history is still advantageous. In addition, we also prevent serious error propagation by periodically flushing the history \mathcal{H} with new training mappings.

B. Classification Algorithms

In this section, we briefly introduce four classifiers we adopt for perimeter distinction. These algorithms fall into two categories.

1) *Discriminant Analysis (DA)*: DA is a suite of fundamental classification algorithms. It is based on models for the class densities. We mainly consider three discriminant analysis methods. (i) Linear discriminant analysis (LDA) fits a multivariate normal density to each class and assumes a common class covariance. (ii) Quadratic discriminant analysis (QDA) also uses multivariate normal density but does not assume the common covariance. It uses a quadratic equation, instead of a linear combination, to describe the boundary between each pair of classes. (iii) The last discriminant analysis method uses the Mahalanobis distance (MDA). It categorizes the input observation by finding the class with the closest centroid using the distance metric [7]. We use discriminant analysis methods because they are simple, diverse, and offer natural low-dimensional views of data [4].

2) *Support Vector Machine (SVM)*: SVM is a more sophisticated classifier. It handles the cases where the classes may not be linearly separable. It generates nonlinear boundaries by building a linear boundary in a large, transformed feature space using kernel functions [4]. We use SVM with a Gaussian kernel to figure out the best performance we can achieve for

given perimeters. We also use the five-fold cross validation to choose the hyper parameters of the algorithm.

C. Measurement Types

We are interested in distinguishing perimeters according to their varied wireless fading conditions. We need measurements that assess fading conditions. RSS and WLS fall into this category.

1) *Received Signal Strength (RSS)*: RSS measures the power present in a received signal. Since wireless fading can manifest itself as either amplification or attenuation of the signal power seen at the receiver, RSS changes are important indicators of fading conditions. Moreover, given that RSS has been extensively used in location estimation (e.g., [8], [9]) and perimeter distinction can be deemed as a coarse granular location estimation, it is tempting to check how well we can perform perimeter distinction using RSS measurements.

2) *Wireless Link Signature (WLS)*: Due to the presence of reflectors in the surrounding environment, the radio waves travel in multiple paths from the transmitter to the receiver when they are reflected, diffracted and scattered in between. Given that the multipath characteristics of the link vary considerably at different locations several wavelengths away from each other, their measurements are a good “signature” of the link. This link signatures is formally defined as [1], [2], [10], [11]:

$$\mathbf{h}_{i,j}^{(n)} = [h_{i,j}^{(n)}(0), \dots, h_{i,j}^{(n)}(\kappa T_r)]^T, \quad (1)$$

where T_r is the sampling interval and $\kappa + 1$ is the number of samples. $h_{i,j}^{(n)}(\tau)$ is the n th channel impulse response measured from the transmitter i to the receiver j at different time τ . As defined in Eq. 2 it is the sum of L paths in the link. For each path l , τ_l is the delay, α_l is the power gain, and ϕ_l is the phase shift.

$$h(\tau) = \sum_{l=1}^L \alpha_l e^{j\phi_l} \delta(\tau - \tau_l) \quad (2)$$

WLS consists of power gains at multiple propagation paths¹, so it not only captures the multipath characteristics of the link but also contains the power information. We explore the richer information in WLS to see if it has a performance advantage over RSS in perimeter distinction.

IV. APPLYING WIRELESS MEASUREMENTS

In this section, we investigate several important issues when applying wireless measurements to perimeter distinction. The most important issue is to determine how capable the wireless measurements are in differentiating fading conditions of various perimeters. For our advanced model, an additional issue is to find out how much loss of this differentiating capability we incur when we have to disregard the power received in the wireless channel. The examination of this issue reveals that not all information in wireless measurements are

¹We do not consider phases of multipath in this work, because we find phases cannot be distinguished across perimeters.

helpful for perimeter distinction. Some information can even be harmful. Therefore, we will look into approaches to condense measurements to improve distinction performance. We also discuss the issue with receiver placement to enable better measurement setup for perimeter distinction. Additionally, we train our perimeter distinction method with multiple measurements within the same perimeter to boost the performance of distinction further. Finally, we use the knowledge gained in analyzing the above issues into our framework to make our perimeter distinction more effective.

A. Distinction Capability When Using Signal Power

In order to understand the distinction capability of wireless measurements, we form the distinction problem in our basic model as a Behrens-Fisher problem [12]. We assume the wireless measurements of different perimeters are normally distributed with unequal means. We perform Student's t-test [13] of the hypothesis that two independent samples $\mathcal{D}_{g1} = \{\mathbf{h}_{g1}^{(i)}\}_{i=1}^{N_{g1}}$ and $\mathcal{D}_{g2} = \{\mathbf{h}_{g2}^{(i)}\}_{i=1}^{N_{g2}}$, from two perimeters $g1$ and $g2$, come from distributions with equal means. The hypotheses are:

$$H_0 : \quad \mathbf{E}(\mathcal{D}_{g1}) = \mathbf{E}(\mathcal{D}_{g2}) \quad (3)$$

$$H_1 : \quad \mathbf{E}(\mathcal{D}_{g1}) \neq \mathbf{E}(\mathcal{D}_{g2}) \quad (4)$$

If the null hypothesis is rejected at a significance level, we conclude that the two samples are from different distributions and thus from different perimeters. Given that the p-value of the test indicates the probability of observing the null hypothesis, small p-values reject the null hypothesis and thus differentiates the two samples. Therefore, the p-values are a good indicator of distinction capability.

Using measurements acquired in a small office and a large office (described later in Section V-A), we apply the t-test on each feature (element) of the WLS vector and get p-values of all features. The empirical cumulative distribution (ECDF) of these values are drawn as the solid line in Fig. 4. 95% of the values are very close to zero and all of them are smaller than the 5% significance level, suggesting a good distinction capability. We also apply the t-test on RSS measurements. This time we only get a single p-value because RSS measurements are scalar quantities. Since this p-value is also very close to zero, RSS has good distinction capability as well. 72% of the WLS p-values are smaller than the RSS p-value, which suggests the best features of WLS can separate fading conditions more effectively than RSS according to the measurements.

B. Distinction Capability Loss When Ignoring Signal Power

Now that we have the assurance that the wireless measurements we choose are capable of distinguishing fading conditions of different perimeters, we next want to know how much we would lose if we cannot rely on the power received in wireless channels for distinction. This is a challenge we must face according to our advanced model. An immediate conclusion we can draw is that the RSS measurements are rendered completely useless under the advanced model because RSS only has power information. WLS has both power and

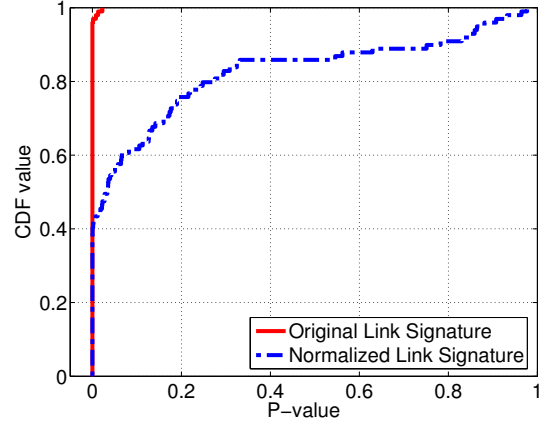


Fig. 4. Empirical CDF of p-values of the t-test on features of the original and normalized WLS using measurements of a small office and a large office.

multipath information. To get rid of the power information, we divide WLS with its magnitude and get the unit-norm signature (*normalized WLS*).

$$\tilde{\mathbf{h}}_{i,j}^{(n)} = \frac{\mathbf{h}_{i,j}^{(n)}}{\|\mathbf{h}_{i,j}^{(n)}\|} \quad (5)$$

where $\|\cdot\|$ is the Euclidean norm. We apply the t-test on the normalized WLS measurements to get a vector of new p-values. We plot the ECDF of the new values as the dot-dashed line in Fig. 4. The figure illustrates a great change in the p-value distribution. There are only a little more than 40% of the new values that are close to zero and only 55% of them are smaller than the 5% significance level. According to the percentage change of p-values within the significant level, WLS loses 45% distinction capable features due to disregarding power.

Feature Selection and Extraction: The analysis in the previous paragraph suggests that the normalized WLS only has a subset of features that have a strong discrimination power while other features are either weak discriminators or bad discriminators (such as those with p-values larger than 0.5). This situation is understandable because while some features are intrinsic to wireless fading conditions, some can be brought about by the measuring instruments, and others may be present due to noise and interference. Therefore, we must retain or magnify the discriminative features and filter out or minimize other features in normalized WLS measurements for higher accuracy.

One approach to achieve this objective is to adopt generic feature selection methods in statistical learning [4]. In Fig. 5 we show the missed distinction rates of feature sets selected by two of such methods. In the first method, features are selected according to their individual p-values in the t-test of the hypothesis defined in Section IV-A. Those with p-values larger than a given significance level are filtered out. In the second approach, features are first added gradually to form different feature sets. Then, each feature set is evaluated with the ten-fold cross validation using a classification algorithm.

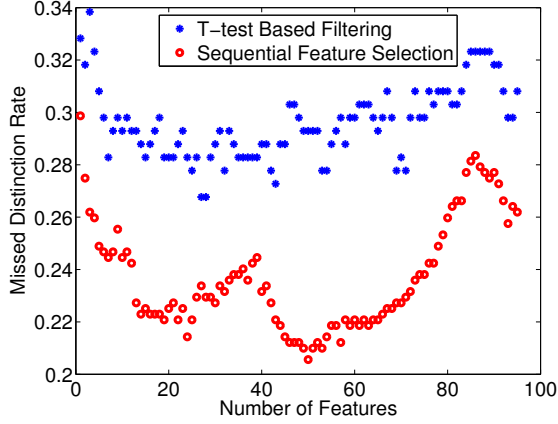


Fig. 5. Performance of features selected by two statistical learning methods.

Finally, feature sets are chosen according to their performance as reported by the algorithm. We use the optimal feature set selected by the second method for perimeter distinction because the first method considers each feature separately without concerning relationships among features. Some features may perform better when they are combined. Thus, features selected by the first method may not be as good as those chose by the second method. As shown in the figure, the missed distinction rates of both methods drop with increasing feature number initially and then fluctuate up and down, but overall the performance of the second method is better. The missed distinction rates in this figure are computed using measurements taken in the small office and the large office.

Another recourse to condense information in the normalized WLS is to extract useful information from features using domain knowledge. We know that the most distinguishing information in the normalized WLS is one that characterizes the multipath channel, so we extract this information by employing metrics that quantify the time dispersive properties of the multipath channel. The most widely used metrics are *Mean Excess Delay* ($\bar{\tau}$) and *RMS Delay Spread* (σ_τ). These are defined as [14]:

$$\bar{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k a_k^2} \quad (6)$$

$$\sigma_\tau = \sqrt{\bar{\tau}^2 - (\bar{\tau})^2}, \text{ where } \bar{\tau}^2 = \frac{\sum_k a_k^2 \tau_k^2}{\sum_k a_k^2} \quad (7)$$

a_k is the magnitude of k_{th} feature in WLS. τ_k is the arrival time of k_{th} feature that is usually measured together with WLS. It is worth noting that the magnitude normalization would not change the values of these two metrics because the ratio of each magnitude to the sum of magnitudes does not change. We utilize both metrics for perimeter distinction not only because they retain the multipath information in WLS but also because they are scalars and, compared with feature sets selected with generic methods, they save more disk space, memory, and computation time.

C. Receiver Placement

Since we distinguish perimeters based on their channel conditions, setting receiver placement to maximize the difference between channels is useful to increase distinction performance. Consider two neighboring perimeters, one cluttered with many scatterers, and one without, and one transmitter placed in each perimeter. If the receiver is placed in the cluttered perimeter, both channels will experience significant multipath fading. If the receiver is placed in the uncluttered perimeter, the channel from the transmitter in the cluttered area will still have significant multipath fading, but the other channel, which has both devices in an uncluttered environment, will have few multipath. In this latter condition, the difference between the two channels is greatest, and will be more easily distinguished. We should thus place receivers in perimeters with fewer scatterers to increase the distinction performance.

D. Burst Based Optimization

In this work we mostly discriminate perimeters created with exterior walls and interior walls, which are arguably the most common perimeters. However, the perimeter can be made of any material. Some perimeters may not attenuate signal power a lot or may not reflect, diffract, or scatter radio waves greatly. In these cases, the receiver placement would not offer us good leverage in increasing the difference between channels of various perimeters. If the original fading difference among perimeters is also small, the distinction performance is bound to be low.

To cope with these difficulties, we need a better understanding of the behavior of wireless users. If we can make a stronger assumption that a certain number of consecutive measurements are transmitted from the same perimeter, we can greatly improve the distinction performance. We call a burst of consecutive measurements from the same perimeter a “measurement burst” and the method to exploit such bursts to improve distinction accuracy the “burst based optimization”. The underlying intuition of this optimization is as follows: There may not be a clear boundary between measurements of two perimeters in the measurement space. A certain percent of measurements from each perimeter may lie at the boundary. These are the causes of missed detection. By taking more samples from each perimeter, we increase the probability that the majority of samples in each burst are not from the boundary area and thus the whole measurement burst can be collectively discriminated. When this probability becomes higher, the distinction performance gets better.

The burst based optimization works in this manner: Instead of reporting the matched perimeter label for individual measurement, our method finds the majority value of matched perimeter labels of all measurements in the burst and then reports this value as the perimeter label for every measurement in the burst. To ensure the existence of the majority value, the burst size must be an odd number. Fig. 6 shows the improved perimeter distinction rates using the burst based optimization. The accuracy of perimeter distinction increases with burst sizes.

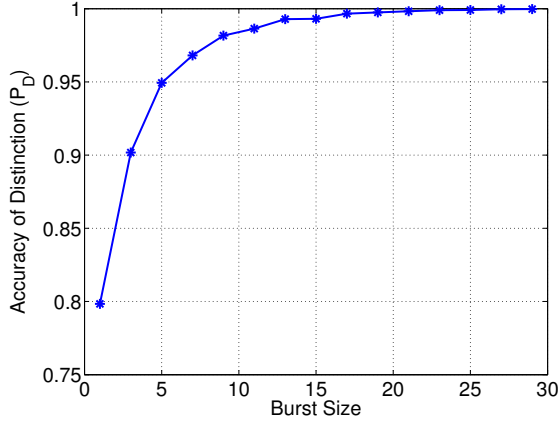


Fig. 6. Accuracy of perimeter distinction under different burst sizes.

E. Advanced Framework

We use the solutions to the issues we discuss in sections IV-A to IV-D to build a more advanced framework for perimeter distinction. In the first step of the new framework, we employ the t-test of the hypothesis defined in Section IV-A to find if all features of input measurements are good discriminators of fading conditions in perimeters. If some features are weak or bad discriminators, we either choose features with sequential feature selection, or extract features using mean excess delay or RMS delay spread. We usually use feature extraction if our main concerns are distinction time, training time, memory and disk usage. Also, we place the receiver in the perimeter with fewest scatters.

Steps two and three are the same as before. In the last step, in our advanced framework, we delay reporting the matched perimeter label for individual measurement and also delay updating the history \mathcal{H} . After we collect matched perimeter labels of all measurements $\{\mathbf{h}^{(N+1)} \dots \mathbf{h}^{(N+m)}\}$ in a burst, where m is the assumed burst size, we find the majority value $g^{majority}$ in the collected label set $\{g^{(N+1)} \dots g^{(N+m)}\}$ and report this value as the perimeter label for each measurement in the burst. Also in this step, the receiver discards the oldest m measurements of the perimeter $g^{majority}$ and adds the mapping pairs $\{\mathbf{h}^{(N+i)}, g^{majority}\}_{i=1}^m$ to the history \mathcal{H} .

V. PERFORMANCE EVALUATION

In this section, we evaluate the performance of our perimeter distinction methodology. First, we describe campaigns we conduct to acquire wireless measurements. Then, we evaluate the basic model by comparing the performance of two measurement types. Finally, we assess the performance of the advanced model from several aspects.

For performance evaluation, we compute the probability of accurate perimeter distinction, P_D , using multiple measurement sets, classification algorithms, and measurement types or features. To evaluate the performance of different methods, we use multi-fold cross validation. We first permute each measurement set, randomly select a subset of measurements to build the training history, and then use the remaining measurements

for testing. We repeat this process for 10,000 times and use the average performance in all tests for evaluation.

A. Measurement Campaigns

We have performed two measurement campaigns. The first one is conducted in several spaces inside a building on a university campus. As shown in Fig. 7, we sample wireless signals when the transmitter moves around in a small office, a large office, an exhibition room, and a hallway. The small office is used by one student for computer vision experiments. The large office is shared by more than ten students as office space. The exhibition room is under construction and has a lot of metal parts when the measurement takes place. The hallway measured in this campaign is outside the exhibition room and the large office. These spaces are separated by interior walls and doors made of wood, plaster, cement, etc. Some spaces are divided by a single perimeter, whereas others are isolated by the combination of perimeters. The receiver is placed at a fixed location in the small office. The measurements are taken when the transmitter moves continuously in this campaign.

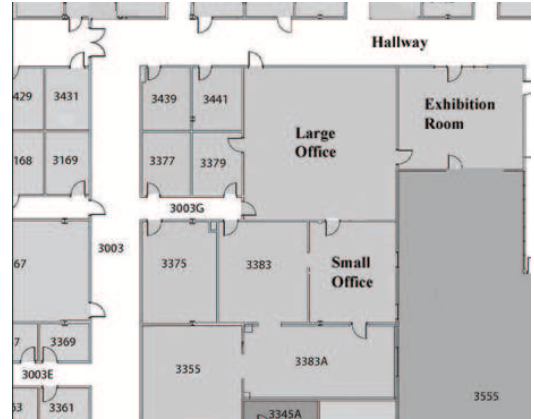


Fig. 7. Floor map of several spaces inside a building.

In the second measurement campaign, we measure wireless signals inside and outside of the same building. The exterior wall of the building is made up of steel and concrete with entrance doors and windows made of aluminum, glass, and plastic. We measure wireless signals along the sidewalk outside of the building and also along the hallway inside the building. Given that we cannot predict how the wireless users move, we measure signals under three different movement profiles in this campaign. In the first profile, we take measurements when the transmitter stops at locations separated by a constant distance along the sidewalk and the hallway. In the second profile, we take measurements when the transmitter stops at random locations on the sidewalk and hallway. In the last profile, we take measurements while the transmitter walks continuously without stopping. The last two profiles offer better realism than the first one, while the first profile allows us to start simple without concerning sudden physical environment changes and changes caused by the Doppler effect [15]. In this campaign, we also study the impact

of the receiver placement on distinguishing indoor wireless measurements from outdoor ones. We put the wireless receiver at a fixed outdoor location first when measuring signals from indoor hallway or outdoor sidewalk. Then, we move it to a mirror location inside the building and perform the same measurement activities again.

We use a pair of Direct Sequence Spread Spectrum (DSSS) transmitter and receiver to measure the attenuation of a test radio signal as well as the occurrence and strength of reflected radio energy known as multipath. The test signal is an unmodulated pseudo-noise code signal with the central frequency of 2443 MHz. It is transmitted with a chip rate of 40 MHz and a code length of 1024.

B. Basic Model Evaluation

WLS vs. RSS: In this subsection, we compare the distinction performance of WLS with RSS. We use four pairs of measurement sets from the first campaign and three pairs from the second campaign for this comparison. The sets from the first campaign contain fading information across four perimeters: the partition between the small office and the large office, one that separates the large office and the hallway, one that divides the large office from the exhibition room, and one that isolates the exhibition room from the hallway. The sets from the second campaign contain indoor and outdoor fading information across a building boundary. They are measured when the transmitter is in different movement profiles as mentioned before. Combining these measurement sets with the four classification algorithms, we form twenty-eight types of evaluations as listed on the x-axis of Fig. 8. These evaluation types are also used in figures 9, 10, and 12.

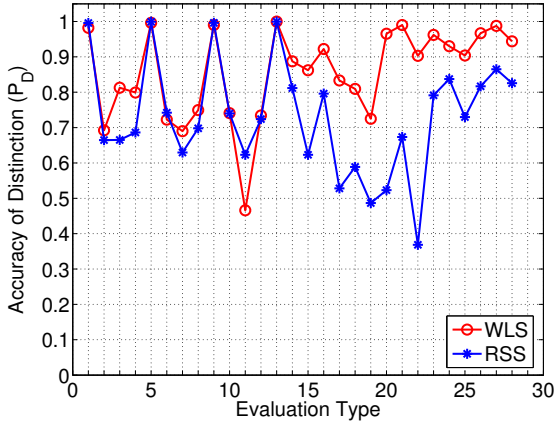


Fig. 8. Comparison of perimeter distinction performance: WLS versus RSS.

Patwari *et al* [1] have shown that for location distinction the RSS-based method has consistently lower performance than the method using WLS. We are interested to know if this conclusion still holds for perimeter distinction. Fig. 8 shows WLS, as denoted by the circle-connected line, outperforms RSS, which is depicted by the star-connected line, in most cases. In twenty-three out of twenty-eight types of evaluations, on average, WLS outperforms RSS by 16.3%. Especially in

the last twelve types of evaluations, which utilize the measurements from the second campaign, WLS shows consistently higher performance than RSS. These results indicate that WLS does have an advantage over RSS for perimeter distinction. This advantage is comprehensible because WLS measurements contain information on multipath propagation channels, which offers additional help in distinction.

However, WLS does not do well in all cases. It has lower distinction rates in five types of evaluations, which all use the measurements from the first campaign. These exceptions can result from several causes. One cause might be that WLS measurements have multiple dimensions so they are harder to categorize, model, and match compared with the one dimensional RSS measurements. Another cause is the limitation of classification algorithms. For example, the low performance of WLS in the 11th evaluation may contribute to the particularity of the MDA.

C. Advanced Model Evaluation

1) *Feature Extraction:* Now we investigate how well the two features that are extracted with the wireless domain knowledge can perform in distinguishing wireless conditions. Fig. 9 shows the distinction rates. The star-connected line shows the rates of mean excess delay, and the circle-connected line indicates the performance of RMS delay spread. When comparing the two lines, it seems mean excess delay has similar performance as RMS delay spread in discriminating fading conditions separated by indoor partitions in the first campaign. The average distinction rate is about 67.9% for mean excess delay and 67.6% for RMS delay spread in these scenarios. However, mean excess delay consistently outperforms RMS delay spread in separating the indoor fading conditions from the outdoor ones across the building boundary in the second campaign. The average rate of mean excess delay is over 83%, whereas the average rate of RMS delay spread is only 77% in these tests. When we compare the first sixteen data points with the last twelve using both features, it is clear that discriminating across the building boundary is much easier than distinguishing across indoor partitions. On average, the distinction rate of indoor partitions is only about 67.7%, but the mean rate of the indoor versus outdoor distinction is more than 80%.

2) *Feature Selection:* We look into the performance of features selected with a generic machine learning method in this subsection. As discussed in Section IV-B, we use sequential feature selection to find the best feature set based on the cross validation results of the training data. To get the optimal distinction performance we individually choose features for each measurement set and for each classification algorithm. We observe that if the classification algorithm used for selecting features is different from the one employed for assessing the selected features, the strength of feature selection is seriously impaired. This situation is true even for the three discriminant analysis algorithms.

The performance of sequential feature selection is depicted by the circle-connected line in Fig. 10. For comparison, we

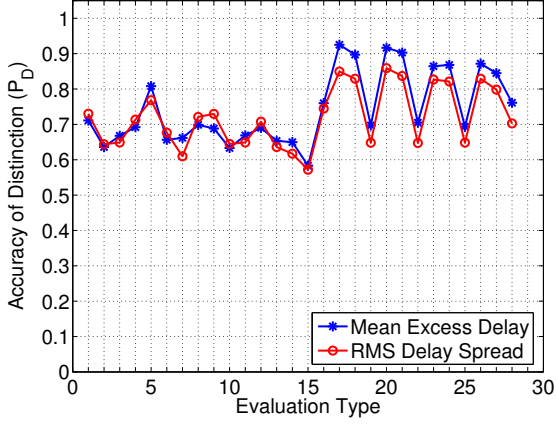


Fig. 9. Distinction using mean excess delay and RMS delay spread.

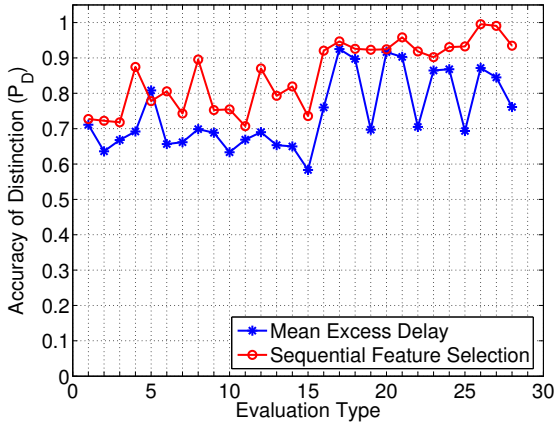


Fig. 10. Perimeter distinction using sequentially selected features.

also plot the distinction rates of the best feature extracted with the domain knowledge, e.g. mean excess delay, using the star-connected line. The selected features outperform the extracted feature in all types of evaluations except the fifth one. The average improvement is about 11% whereas the only decrease is 3%. Although we gain considerable performance boost from sequential feature selection, it takes a pretty long time to find a good feature set. In addition, the knowledge in the wireless domain allow us to summarize 100 raw WLS features into one, while the selected set usually has 30 to 50 features. More features take more memory, disk space, and computation time. More features also requires more measurements for training. For some classification algorithms such as LDA, if the training set size is not large enough for the number of multiple features chosen, overfitting may occur or the resulting low-rank matrices can lead the training process to fail.

3) *Receiver Placement*: In Section IV-C we discussed the issue with receiver placement. We argued that when placing the receiver inside the right perimeter the measurement setup can improve the distinction performance; on the other hand, if we put the receiver in the wrong place, the accuracy of perimeter distinction can be compromised. In this section we utilize wireless measurements from the second campaign

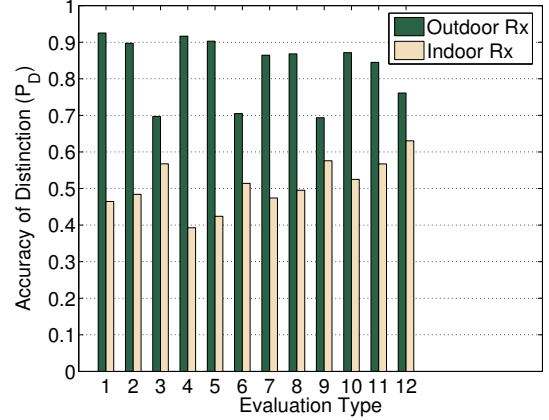


Fig. 11. The performance degradation in distinguishing indoor from outdoor environment when moving the receiver from outdoor to indoor.

to validate and assess the impact of receiver placement on perimeter distinction.

As explained in Section V-A, we rotated the receiver between two mirror locations, one indoor and another outdoor, when measuring wireless signals across the indoor and outdoor perimeter. Fig. 11 shows the performance comparison based on three pairs of measurement sets taken with the indoor receiver and another three pairs measured with the outdoor receiver. Combining the three sets of each receiver location with the four classification algorithms, we form twelve evaluation types in Fig. 11. The performance degradation when moving the receiver from outdoor to indoor is so great that on average there is an 32% drop in distinction rate. This figure shows the results as evaluated using mean excess delay. We also find the similar performance degradation caused by the wrong receiver placement in the evaluation results using RMS delay spread. The indoor receiver only suffers the great performance degradation when it cannot use signal power to make the distinction as required by the advanced application model. In the basic model, the indoor receiver can still achieve the distinction accuracy up to 90%.

4) *Burst Based Optimization*: In Fig. 12, we show the performance improvement by using burst measurements of RMS delay spread. As we expect, the use of burst measurements is very effective and works for all measurement sets and classification algorithms. On average there is a 26% improvement. The final improved mean P_D rate is 97.7%. When applying the burst based optimization to mean excess delay, we similarly obtain great distinction improvement: 24% on average, which makes the final mean P_D rates to reach 98.8%. In this assessment, we use the burst size up to 29. Longer bursts can further improve the perimeter distinction performance, however, collection of longer bursts might not always be possible.

VI. RELATED WORK

Perimeter discovery [16] has been studied in wireless sensor networks. It focuses on how sensor nodes can determine if

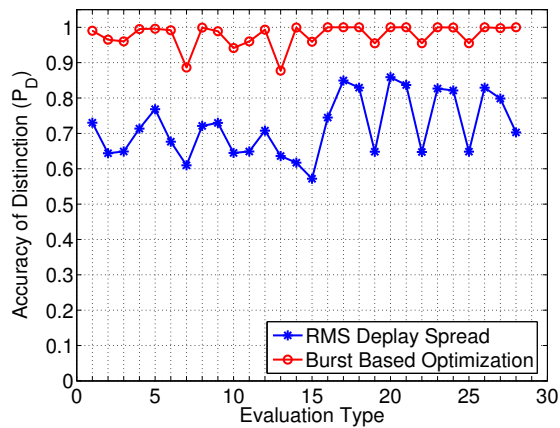


Fig. 12. Improved distinction performance using the burst based optimization with RMS delay spread.

they are located along the perimeter of a sensor network. Our objective is less precise; we only want to know at which side of the perimeter the node resides. On the other hand, our problem is not less challenging, because we do not rely on localization information as perimeter discovery does and we do not require some nodes to be located along the perimeter. Our work also bears some similarity as the Device-free-Passive (DfP) localization in [17]. Both our work and the work of [17] are based on the fact that RF signals are affected by changes in the environment. However, the objectives of the DfP system are detecting, tracking, and identifying entities in an environment, while our objective is to distinguish environments themselves.

There are also some related work in context awareness studies [18], [19], [20]. In [19] the authors propose to use light intensity to distinguish rooms. In [20] a wearable camera and microphone are utilized for recognizing an individual's situation. A recent work in ambience fingerprinting [18] made further advancement in this area. The researchers made full use of the camera, microphone, and accelerator in a mobile phone to create a photo-acoustic plus motion signature and used it to recognize logical locations such as Starbucks and McDonalds. These work have similar objectives as ours, but they mainly leverage ambient sound, light, and motion in a environment, whereas we concentrate on wireless fading conditions that are affected by static environment setup as well as dynamic changes. Furthermore, we propose a single-receiver measurement setup that can increase the fading difference at two sides of a perimeter with the right receiver placement, whereas in [18] the wireless receiver always works within the same perimeter as the transmitter.

VII. CONCLUSIONS AND FUTURE WORK

We used PHY-layer wireless measurements for perimeter distinction. We described two application models and an algorithmic framework for perimeter distinction. We investigated important issues in applying wireless measurement to perimeter distinction, and evaluated the performance of perimeter distinction with two types of measurements and four classification algorithms. Our experimental evaluation

showed that our methodology can achieve accurate perimeter distinction up to 80% even without using signal power and we are able to improve the accuracy up to 98% using burst measurements. In the future, we plan to explore the use of Multiple Signal Classification algorithms [21] to obtain WLS with higher resolution channel responses, and also consider a multi-receiver collaborative framework for further improving perimeter distinction accuracy.

REFERENCES

- [1] N. Patwari and S. K. Kasera, "Robust location distinction using temporal link signatures," in *ACM Intl. Conf. on Mobile Computing Networking (Mobicom '07)*, Sept. 2007.
- [2] J. Zhang, M. H. Firooz, N. Patwari, and S. K. Kasera, "Advancing wireless link signatures for location distinction," in *ACM Intl. Conf. on Mobile Computing Networking (Mobicom '08)*, Sept. 2008.
- [3] E. Elnahrawy, X. Li, and R. Martin, "The limits of localization using signal strength: A comparative study," in *Proceedings of The First IEEE International Conference on Sensor and Ad hoc Communications and Networks (SECON)*, 2004, pp. 406–414.
- [4] T. Hastie, R. Tibshirani, J. Friedman, and J. Franklin, "The elements of statistical learning: data mining, inference and prediction," *The Mathematical Intelligencer*, vol. 27, no. 2, pp. 83–85, 2005.
- [5] H. Daumé, III and D. Marcu, "Domain adaptation for statistical classifiers," *Journal of Artificial Intelligence Research*, vol. 26, no. 1, pp. 101–126, 2006.
- [6] H. Daumé, "Frustratingly easy domain adaptation," in *Annual meeting-association for computational linguistics*, vol. 45, no. 1, 2007, p. 256.
- [7] P. Mahalanobis, "On the generalized distance in statistics," *Proceedings of the National Institute of Science, Calcutta*, vol. 12, 1936, p. 49.
- [8] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *IEEE INFOCOM 2000*, 2000, pp. 775–784.
- [9] N. Patwari and A. O. Hero III, "Using proximity and quantized RSS for sensor localization in wireless networks," in *2nd ACM Workshop on Wireless Sensor Networks & Applications (WSNA'03)*, Sept. 2003, pp. 20–29.
- [10] N. Patwari and S. K. Kasera, "Temporal link signature measurements and models for location distinction," submitted to *IEEE Transactions on Mobile Computing*.
- [11] J. Zhang, S. K. Kasera, and N. Patwari, "Mobility assisted secret key generation using wireless link signatures," in *INFOCOM mini-conference*, 2009.
- [12] M. Fligner and G. Policello, "Robust rank procedures for the Behrens-Fisher problem," *Journal of the American Statistical Association*, pp. 162–168, 1981.
- [13] J. Box, "Guinness, Gosset, Fisher, and small samples," *Statistical Science*, vol. 2, no. 1, pp. 45–52, 1987.
- [14] T. S. Rappaport, *Wireless Communications: Principles and Practice*. New Jersey: Prentice-Hall Inc., 1996.
- [15] K. Toman, "Doppler and the Doppler Effect," 1984.
- [16] A. M. Khedr, W. Osamy, and D. P. Agrawal, "Perimeter discovery in wireless sensor networks," *Journal of Parallel and Distributed Computing*, vol. In Press, Corrected Proof, 2009.
- [17] 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*. ACM, 2007, p. 229.
- [18] M. Azizyan, I. Constandache, and R. Choudhury, "SurroundSense: Mobile Phone Localization via Ambience Fingerprinting," in *ACM MobiCom*, 2009.
- [19] N. Ravi and L. Iftode, "FiatLux: Fingerprinting Rooms Using Light Intensity," in *Adjunct Proceedings of the Fifth International Conference on Pervasive Computing*. Citeseer, 2007.
- [20] B. Clarkson, A. Pentland, and K. Mase, "Recognizing user context via wearable sensors," *Wearable Computers, IEEE International Symposium*, vol. 0, p. 69, 2000.
- [21] K. Pahlavan, X. Li, and J. Makela, "Indoor geolocation science and technology," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 112–118, 2002.