2013 IEEE 33rd International Conference on Distributed Computing Systems

Pilot: Passive Device-free Indoor Localization Using Channel State Information

Jiang Xiao*, Kaishun Wu*†‡, Youwen Yi*, Lu Wang* and Lionel M. Ni*

*Department of Computer Science and Engineering
Guangzhou HKUST Fok Ying Tung Research Institute, HKUST
†National Engineering Research Center of Digital Life
State-Province Joint Laboratory of Digital Home Interactive Applications, Sun Yat-sen University
‡Corresponding Author

***Abstract*—Many emerging applications such as intruder detection and border protection drive the fast increasing development of device-free passive (DfP) localization techniques. In this paper, we present *Pilot*, a Channel State Information (CSI)-based DfP indoor localization system in WLAN. *Pilot* design is motivated by the observations that PHY layer CSI is capable of capturing the environment variance due to frequency diversity of wideband channel, such that the position where the entity located can be uniquely identified by monitoring the CSI feature pattern shift. Therefore, a "passive" radio map is constructed as prerequisite which include fingerprints for entity located in some crucial reference positions, as well as clear environment. Unlike device-based approaches that directly percepts the current state of entities, the first challenge for DfP localization is to detect their appearance in the area of interest. To this end, we design an essential anomaly detection block as the localization trigger relying on the CSI feature shift when entity emerges. Afterwards, a probabilistic algorithm is proposed to match the abnormal CSI to the fingerprint database to estimate the positions of potential existing entities. Finally, a data fusion block is developed to address the multiple entities localization challenge. We have implemented *Pilot* system with commercial IEEE 802.11n NICs and evaluated the performance in two typical indoor scenarios. It is shown that our *Pilot* system can greatly outperform the corresponding best RSS-based scheme in terms of anomaly detection and localization accuracy.**

***Index Terms*—Device-free Indoor localization, Channel State Information, RSS, Physical Layer.**

I. Introduction

Indoor location based services (LBSs) are becoming ubiquitous popular for providing people location-aware information. Advances have been made to enable the indoor LBS using RF-based technologies such as WLAN, wireless sensors and Radio-frequency identification (RFID), etc. Most of these technologies share a common requirement that special devices like WiFi-enabled smartphones or RFID tags must be carried. However, as LBSs are bringing forth new expectations, such device-based approaches become ineligible for satisfying some emerging application demands. For example, exhibition galleries and shopping centers are expecting to support the pilferage prevention and missing people tracking services in a way that the visitors and customers do not need to carry on specific hardware. Important applications also exist in other indoor settings like hospitals, residences and places of entertainment. In hospital, health care providers need to grasp the distribution of location of the wandered patients associated without a device and quickly expand the relief operations. Also, people can figure out the position of the intrusive individuals in a resident district for safety precaution. Therefore, a device-free passive technique capable of detecting, positioning, and tracking entities neither carry any devices, nor participate actively in the localization process will be greatly helpful.

In general, the underlying technical challenges for designing a passive device-free indoor localization system are in two folds. First noted that device-based class can inherently obtain the knowledge of current status with a device attaching to the target and directly do localization. In contrast, the nature of device-free scheme requires implementing a similar functionality by detecting the occurrence of anomalous entity in the area of interest. Therefore, the first challenge that has to be addressed in order to enable the novel location-aware applications is anomaly detection problem, also known as human/motion detection. Second, how to accomplish localization when a motion event of an entity has been detected serves as the new knotty problem. State-of-the-art researches [3], [8], [9] adopt radio signal strength (RSS) as the base modality in an attempt to overcome these difficulties. However, we argue that the performance of device-free positioning systems based on RSS is limited by the disadvantage of RSS itself. Specifically, indispensable anomaly detection can be suffered from the high variability of RSS, owing to its coarse measurement. Moreover, the inherent fluctuation of RSS makes it less sensitive to entity-caused environmental changes, not to precisely signify a location fingerprint. Consequently, there is a pressing need to prompt a new modality superior to RSS for device-free indoor localization.

Fortunately, physical layer Channel State Information (CSI) from OFDM-based system promises new potential to overcome the above limitations of RSS. Previously, CSI has proved to be a reliable metric for locating the entity with WiFi-enabled device [24]. Under this ground, we envision a future of leveraging CSI for passive device-free indoor localization. To this end, we start by investigating the feasibility of CSI-based device-free scheme. Based on preliminary experiments, we obtain two key observations. The primary one is that CSI

1063-6927/13 $26.00 © 2013 IEEE
DOI 10.1109/ICDCS.2013.49

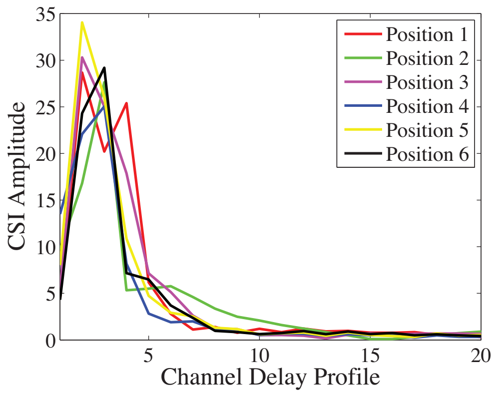


Fig. 1: Delay Profile in Different Environments.

is capable of detecting anomaly that affected by changes in the environment. This relies on the temporal stable feature of CSI that ensures the sensitivity of capturing the environmental variance owing to abnormal entities' (i.e., human) occurrence and movement. The second insight stems from adequacy of CSI to differ a fixed location where the entity is present from all the other locations. Frequency diversity [11] of CSI allows it to reflect the varying multipath reflections due to entities' existence. In Figure 1, the time domain delay profile obtained by inverse fast Fourier Transform (IFFT) of frequency domain CSI shows that an entity in different positions will change the multipath reflections differently and result in different delay profiles. Thus, CSI offers two major benefits when detecting the abnormal entities and serving as location fingerprint.

In this paper, we design **Pilot**, a CSI-based **P**assive device-free **i**ndoor **lo**calizat**t**ion system. Our main idea is to leverage the beneficial characteristics of CSI to monitor the abnormal appearance (anomaly or motion detection) and then to identify the location of entity. In particular, we design three blocks to enable the passive device-free localization functionality. First, we explore the frequency diversity of CSI in passive radio map construction block to generate normal and abnormal fingerprints. Second, anomaly detection block utilizes the correlation of CSI over time to monitor the abnormal variance. This block is the prerequisite of finalizing locating the position of the anomaly entities, which is more challenge. Third, we tackle this knotty localization problem with position estimation block. Nevertheless, we develop a data fusion algorithm to determine the positions of multiple entities.

The main contributions of *Pilot* system are summarized as follows:

1) We exploit the feasibility of using fine grained channel state information for passive indoor localization. To the best of our knowledge, this is the first work to leverage PHY layer information CSI for DfP indoor localization in WLAN.
2) We take the advantages (temporal stability and frequency diversity) of CSI to design *Pilot*, a passive indoor localization system, to realize passive radio map construction, anomaly detection, and position estimation, respectively.
3) Extensive evaluations of *Pilot* with commercial 802.11 NICs are conducted in two typical indoor scenarios. These measurements show that the *Pilot* provides higher anomaly detection ratio than RSS-based RASID. *Pilot*, Pilot greatly outperforms RSS-based Nuzzer system with respective to localization accuracy.

The remainder of this paper is structured as follows. Section III discusses the central two observations that motivate our approach. Then we summarize the state-of-the-art researches on indoor localization in Section II. Section IV presents the overall architecture design of *Pilot* along with detailed methodology. In Section V, we describe the implementation of Pilot, and evaluate the performance in two typical indoor environments. Finally, we render our conclusions and present avenues for further research based on this work in Section VI.

II. Related Work

Indoor localization has gained worldwide attention for its advantages of providing location awareness for various kinds of LBSs. There are primarily two categories of techniques related to this and become an increasing popular research field, namely *device-based* and *device-free* techniques. We expand upon representative prior studies in each of these two-fold techniques below.

Device-based techniques: Existing and emerging indoor localization systems mainly depends on device-based techniques that targeted entities can only be localized with attaching a device. To name a few, LANDMARC [19] employ densely deployed RFID tags as receiver in the positioning region of interest; Cricket [27] and Active Badge [16] separately handle the localization problem by leveraging ultrasonic and infrared sensors; wireless sensors such as MicaZ [7] and TelosB [10] employed in various scale testbed enable an alternative approach for location estimation; FM radio [17] is also proposed for positioning purpose. However, they all require a specific hardware to facilitate measurements for localization. In addition, some of them constrained to particular conditions such as infrared can function with the necessity of light of sight (LOS) existence. Alternatively, the pioneer RADAR [28] system investigates radio signal strength (RSS) to measure the distance between the APs and WiFi-enable receivers. Horus [20] improves the accuracy by applying a probabilistic model of RSS distribution. To avoid time-consuming site survey, WILL [23] augments user motions with the RF signal characterisers to construct a logical radio map for localization. In [4], [5], [6], the authors propose to use temporal channel response as link signature for differentiating locations. Even though these systems employ the already installed WLAN infrastructure without additional cost, they still require efforts from carrying on device at the transmitter that inappropriate for ubiquitous scale setting. On the contrary, our Pilot system is purely passive and device-free.

Device-free techniques: Driven by the necessity of satisfying expectations of new kinds of location-aware services, device-free techniques [1] - do not require the entities to carry any device - have gained widespread concern by research community.

Computer vision [18], [22] and RFID tags [9], [31] have been deployed in an indoor environment for device-free localization functionalities. In a similar fashion, the authors propose a similar concept of "transceiver-free" [8], and use wireless sensor networks for building a RF-based object tracking system [21], [9]. In [29], [22], Radio Tomographic Imaging (RTI) technique is presented for imaging the passive moving objects by applying a linear model. In [26], the authors makes improvement by leveraging motion-induced variance of RSS measurements. Yang et. al. [12] develop a joint learning GREEK algorithm to effectively diagnose the presence of intrusions in Zigbee network. However, the above approaches loss attraction in terms of scalability due to either high specific hardware cost like video camera and RFID tags and maintainable cost (i.e., sensors). WLAN-based approaches [1], on the other hand, use the available infrastructure for indoor localization. Recently, the authors have developed a RSS-based Nuzzer system in a large scale indoor setting with pre-installed WLAN infrastructure [3]. In this paper, we introduce the use of a new metric CSI from PHY layer for device-free indoor localization to replace the coarse RSS value, which can be resist to temporal variance and sensitive to environment changes by exploiting the frequency diversity. To the best of our knowledge, this is the first work to apply fine-grained CSI to improve performance of device-free indoor localization in WLAN.

III. Background and Hypotheses

A. PHY Layer Channel State Information

Our system leverages CSI value for device-free indoor localization. We therefore review such CSI value in this section.

In wireless communications, Channel State Information (CSI) is a fine-grained PHY layer information that describes the channel property of a radio frequency (RF) link at the subcarrier level. To be more specifically, CSI describes how a RF signal propagates from the transmitter(s) to the receiver(s) and reveals the combined effect of, for instance, scattering, fading, and power decay with distance. Generally, CSI is a collection of $M \times 1$ matrices H that specifies channel gain over a pairs of transmitter and receiver with multiple antennas over M subcarriers. Mathematically, CSI on a single subcarrier can be represented by amplitude($|h|$) and phase($\angle h$) as $h = |h|e^{j\sin\{\angle h\}}$. Based on IEEE 802.11n standard, the commercial wireless network interface card (NIC) allows us to obtain CSIs conveniently.

CSI based on OFDM system has gained popularity in a couple of applications. To name a few, authours in [30] propose to utilize CSI for rate adaptation instead of widely used RSS. That is, wireless packet delivery can be accurately predicted by using CSI. In [24], CSI is shown to be appropriate for device-based indoor localization. CSI offers the capability of estimating the distances between transmitters and receivers. In [25], we introduce the use of CSI for indoor motion detection, which can be sensitive to environment changes and resist to temporal variance. In this paper, we further explore the favorable features of CSI for realizing device-free indoor positioning.

B. Hypotheses and Measurements

In this section, we start our work by testing two hypotheses of utilizing CSI for device-free localization. We demonstrate that, such hypotheses provide insight for our eventual system design. Afterwards, based on preliminary measurements, we validate these hypotheses and shed some light on the design of a new CSI-based device-free localization system.

We present two integrant hypotheses of designing a CSI-based device-free localization system as follows:

Hypothesis 1: CSI over multiple subcarriers can reveal the abnormal status caused by appearance of human. More specifically, a motion behavior of an entity will cause CSIs variance that exhibits some kind of feature pattern shift.

To support typical device-free location-aware applications such as intruder localization, the primitive step we need to conduct is detecting the presence of a suspicious device-free entity, i.e., the motion. This motion behavior indicates an abnormal event happened during the whole localization process, termed as a "localization trigger". Once a "trigger" is detected, the location estimation block is followed and finalized. However, how to perform passive anomaly detection to serve as "localization trigger" is challenging as the intrusive entities usually do not carry any radios or may not even cooperative. Recently, RSS has been used for device-free motion detection in RASID system [2]. However, the fatal flaw of RSS lies in its susceptibility to measurement itself due to severe multipath effect in indoor environment. For this reason, we consider to study the feasibility of leveraging a more reliable CSI for detecting the mobility of entity. The intuition is to investigate whether the CSIs over multiple links can appropriately infer a moving entity by extracting the feature patterns. In our design, we first attempt to apply single pair of AP and DP to monitor this "trigger" occurrence.

Hypothesis 2: CSI over multiple subcarriers can be leveraged to distinguish entity, i.e., human, in different locations. That is, CSIs show some kind of differential feature patterns when an entity appears at different locations.

After a motion behavior is detected, next comes to identify the location of the entity in the region of interest. In fact, the presence of an entity will influence the RF links between APs and DPs in a typical indoor environments. On the basis of the location of entity, the two-way impacts can be classified as:

- *Direct light-of-sight (LOS) blocking*: the entity is located exactly between the AP and DP such that blocks the direct LOS transmitting link;
- *Indirect non-light-of-sight (NLOS) reflection*: the position of entity lies beside the LOS link and influences the multipath propagation of RF signals.

CSI is capable of revealing the change of channel status due to the blockage of LOS path. In addition, it can present multiple NLOS reflection due to frequency diversity as shown

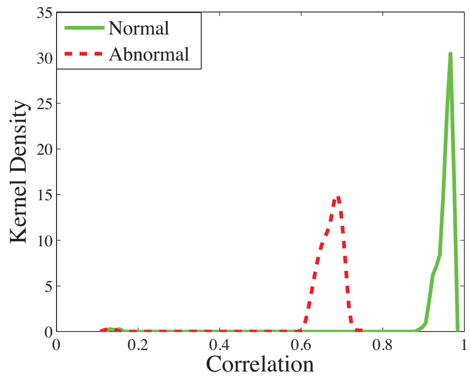


Fig. 2: Anomaly Detection by CSI Feature Shift.

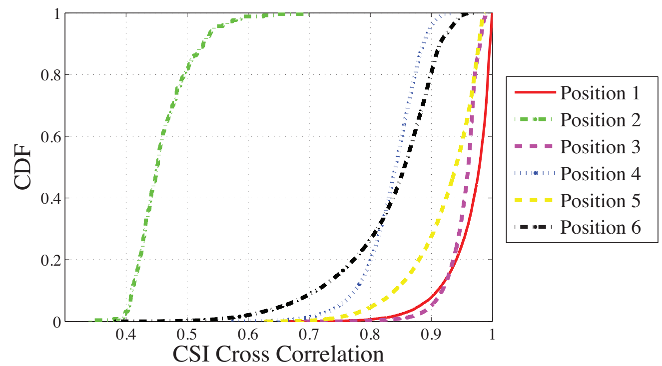


Fig. 3: Location Distinction by Variant CSI Features.

in Figure 1. Therefore, we hypothesize that the CSI will exhibit unique feature pattern at a given location that different from the others due to the impact of an entity's motion behavior. In this way, a "passive" radio map can be constructed by storing the CSI over each RF link as a fingerprint for each location.

We verify these two hypotheses using the following preliminary experiments.

Experiment 1: Anomaly Detection by CSI Feature Shift

Our first experiment examines the effects on CSI when an anomalous entity is appeared in the monitoring region. We expect that CSI will exhibit distinguishable characteristics between static status and dynamic status. Under static status, we collect CSIs of n packets and store them into a passive radio map, namely normal fingerprints database (DB) $\mathbb{H}_{Nor}$ as:

$$\mathbb{H}_{Nor} = (H_1, H_2, \ldots, H_n) \tag{1}$$

We calculate the self-correlation of the CSIs of each packet i with all the rest packets and afterwards average the aggregated correlation sum as follow,

$$\mathbb{C}^i_{Nor} = \frac{1}{n}\sum_{j=1}^{n} corr(\mathrm{H}^{\mathrm{i}}_{\mathrm{Nor}}, \mathrm{H}^{\mathrm{j}}_{\mathrm{Nor}}) \tag{2}$$

This $\mathbb{C}^i_{Nor}$ is set as a "normal" correlation value for detecting an abnormal behavior. We set up an abnormal environment where only one person is present in the area of interest. Similarly, we measure the CSIs and compare them to the constructed normal profiles by applying correlation function as:

$$\mathbb{C}^i_{Abn} = \frac{1}{n}\sum_{j=1}^{n} corr(\mathrm{H}^{\mathrm{j}}_{\mathrm{Nor}}, \mathrm{H}^{\mathrm{i}}_{\mathrm{Abn}}) \tag{3}$$

Figure 2 plots the empirical probability density function (PDF) of CSIs feature patterns between these two kinds of statuses. Clearly, there is an obvious feature shift of CSI correlation when encountering an anomalous event. It means that CSI is temporal stable in static environment while sensitive enough for an instantaneous motion response.

Experiment 2: Location Distinction by Variant CSI Features

Our next experiment is to inspect if the variant feature patterns of CSIs show uniqueness for a given location with entity and can be applied as fingerprints for localization.

In Figure 3, we depict the cumulative distributive function (CDF) of CSI feature pattern across 6 locations. To be specific, the red curve presents the CDF of self-correlation for position 1, while other 5 curves show the CDF of cross-correlation between each with position 1, respectively. It is observed that one location can be distinguished from the others by analyzing the statistical properties of cross correlation. More concretely, these CSI feature patterns provide appropriate information regarding differentiating locations as fingerprints.

In summary, we have made two important observations in this section:

1) CSI can capture the environment variance due to its temporal stability;
2) CSI can differ a given location where the entity appears from all the other locations.

This motivates us to apply CSI in device-free technique to achieve high localization accuracy. In what follows, we detailed design a novel CSI-based passive indoor localization system.

IV. The Pilot Design

In this section, we present the design of Pilot system. We begin with an overview of Pilot architecture with three key constituent blocks. Then, we lay out detailed description of each block in subsequent subsections.

A. Overview

Pilot is built on the WLAN infrastructure without additional deployment and management cost. In our design, Pilot consists of three hardware elements: access points (APs), detecting points (DPs), and Pilot server as depicted in Figure 4. A radio frequency (RF) link will be established between a pair of AP and DP kept stationary during the whole localization period. The AP broadcasts beacon message periodically. Besides involved in the activity of localization, those APs can also serve as hotspots simultaneously. The DP is a general wifi compatible device which is responsible for interacting with

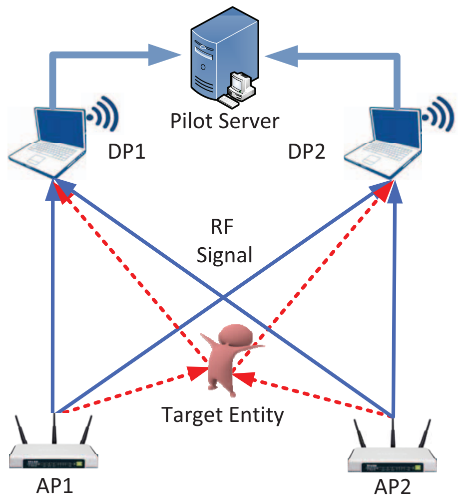


Fig. 4: Hardware elements of Pilot.

both AP and server. Upon receiving beacon messages from the AP, the DP will record the according CSI. These raw CSIs across multiple subcarriers at the PHY layer are then uploaded to the Pilot server. Figure 5 shows the architecture of Pilot system. Pilot server will perform localization by carrying out three main procedures as below.

1) ***Passive Radio Map Construction*** Passive Radio Map Construction block is primarily developed for two purposes: 1) to generate a normal CSI profile for anomaly detection; 2) to build up an abnormal database corresponding to different positions where entity is located for position estimation. It is worth mentioning that this radio map is "passive" since its generation involves no active participation with device-based entity.
2) ***Anomaly Detection*** We then design the Anomaly Detection block to capture environmental changes due to entity's appearance. Pilot continuously executes monitoring of CSI feature variance that indicates whether an entity emerges or not in the area of interest. Once an abnormal event has been detected, it triggers the immediately subsequent execution of localization.
3) ***Position Estimation*** The Position Estimation block is designed to map the instant abnormal CSIs to passive radio map, and fuse the data over multiple links. Such that the exact location of an abnormal entity can come to knowledge.

In the following subsections, we will describe each block of Pilot in a divide-and-conquer manner.

B. Passive Radio Map Construction

First, we propose an offline block - Passive Radio Map Construction - as a basis to facilitate subsequent operations in fingerprinting system Pilot. Map construction block consists of two functions: processing the measurement data and generate the fingerprints database. In OFDM-based networks, a RF signal is transmitted over multiple subcarriers simultaneously occupied a wide band. In a typical indoor setting, the channel is affected by frequency-selective fading and frequency independent attenuation. The coarse RSS fails to uncover the channel state at subcarrier level. In comparison, significant diversity over heterogeneous subcarriers can be captured by CSI. Therefore, the underlying idea is to exploit the frequency diversity of CSI, which reveals diverse feature patterns of different locations due to appearance and movement of abnormal entities.

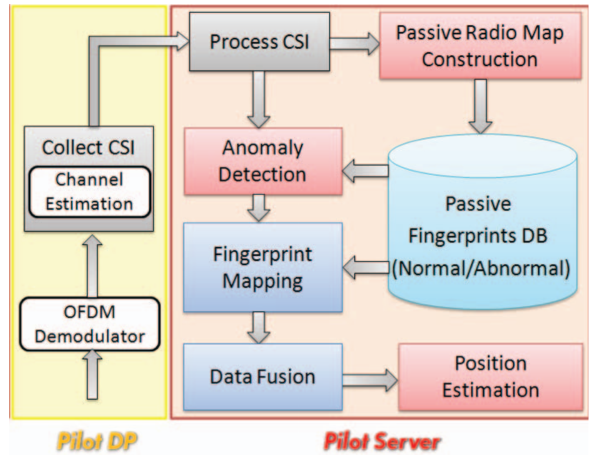


Fig. 5: Pilot Architecture.

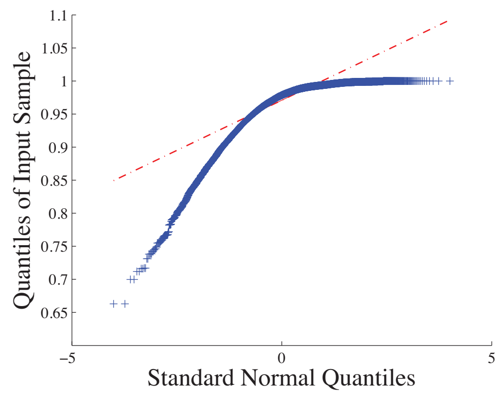


Fig. 6: CSI correlation Gaussian Hypothesis test.

To begin with, we collect CSIs samples over the RF links between APs and DPs in a normal state, i.e., with no entity appearance. We then process these sample data received from multiple DPs on Pilot server. A normal fingerprints database denoted as $\mathbb{C}_{Nor}$ composes of a set of n processed CSI samples will be constructed. As stated in previous section, the process of the normal profile construction is to average the sum of the self-correlation of CSI samples over each RF link. We use $QQ-plot$ to test the Gaussian distribution hypothesis of $\mathbb{C}_{Nor}$, and Figure 6 shows that it doesn't pass Gaussian hypothesis test, nor tests for other well-known distributions.

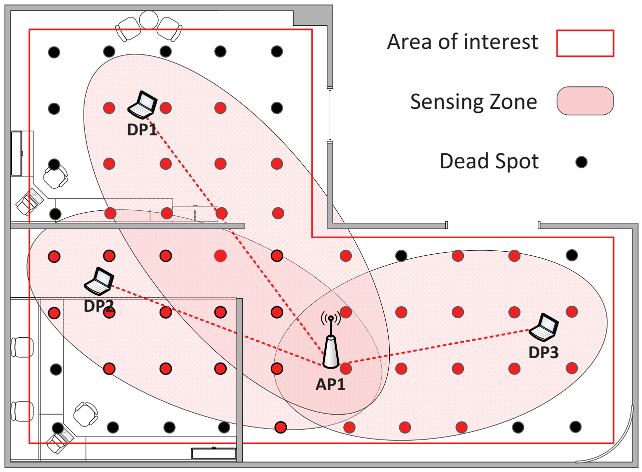


Fig. 7: Sensing Zone and Dead Spot.

Note that the calibrated normal profile is a premise for anomaly detection. In what follows, we need to generate a passive radio map in an offline phase for indoor localization. Unlike map construction of device-based fingerprinting system, this passive radio map construction is conducted with entities located at different positions without any devices, and the CSIs are collected by the DPs over each RF link. Similarly, the self-correlation of the CSIs for each position is stored in the database.

With the objective of narrowing the fingerprint mapping range, and also improving the localization accuracy with single link, we introduce two novel concepts as following:

Definition 1. ***Sensing Zone*** *is the zone where the present of entity can be detected by a specific AP-DP link.*

Definition 2. ***Dead Spot*** *is the position in the area of interest that outside the sensing zone of any AP-DP link.*

Intuitively, the further the entities away from the monitor link, the less influence its appearance on the CSI measurements. Hereby, "*Sensing Zone*" is christened to capture the influence of abnormal entities on RF links as shown in Figure 7. If the entity is positioned in the "*Dead Spot*" (i.e., outside the sensing zone), the server will not be able to distinguish this situation between normal one, and results in false negative.

Note that, due to the multipath reflection and LOS blocking of signal in indoor environment, the shape of the Sensing Zone of a transceiver link is not a regular one like circle or ellipse, but much more complicated. Therefore, the Sensing Zone can only be characterized with field measurements. In this work, we determine the Sensing Zone with CSI correlation feature for each link as following. We calculate the cross correlation of CSIs on each position in the passive radio map and the normal CSIs for each link according to Eq.(3). Then we will check the density distributions of $\mathbb{C}_{Nor}$ and $\mathbb{C}_{Abn}$ are significantly different. If the difference is big enough, the position is within the Sensing Zone of this RF link. Since the distribution is not Gaussian or well-approximated by other known distributions, we can use nonparametric hypothesis test. Specifically, we perform *Ansari-Bradley test* [13] where the null hypothesis defined as:

H_0: $\mathbb{C}_{Nor}$ and $\mathbb{C}_{Abn}$ follow the same distribution.

If H_0 of identical distributions cannot be rejected at the $\alpha\%$ significance level, we can conclude that they follow the same distribution, and the corresponding position is outside the sensing zone of the specific link. With these two concepts, we can reduce the searching space when motion is detected by a specific RF link. In addition, we can also optimize the positions of transmitters and receivers, so that all the monitor spots in the area of interest can be covered with minimal deployment cost.

C. Anomaly Detection

A process of determining abnormal events from CSI measurements is a prerequisite for device-free indoor localization. Anomaly Detection block is designated to continuously execute this process. As mentioned in Section III, we observe that CSI stays relatively stable over time in a normal state (i.e., static) without entities' appearance or movement. Whereas in an abnormal status (i.e., mobile), CSI experiences a feature pattern shift. Therefore, the basic idea is to leverage the temporal stability characteristic of CSI consistent with normal status so as to distinguish from the feature patterns under abnormal environments. In Pilot, we utilize this different feature pattern shift as a "localization trigger" for deciding whether a localization process should be started.

The main idea is to check the probability of each CSI feature to be in normal profile. To decrease the effect of outliers, we use a sliding window to average those raw CSI measurements. Now we attempt to estimate the probability of the current CSI correlation sample $\mathbb{C}$ according to the statistics of $\mathbb{C}_{Nor}$. However, the distribution of normal profile cannot be assumed as Gaussian and an alternative method for distribution estimation is in need. Therefore, we alternatively adopt a kernel density-based function approach as explained below.

In statistics, this approach is known as kernel density estimation (KDE). The benefit of KDE is that it can estimate the density directly from the data without assuming a particular form for the underlying distribution. For RF link l between a pair of AP and DP, we denote W as the sliding window length, and n as the total number of correlation samples in the normal profile. Now consider a sequence of independent and identically distributed (i.i.d.) random correlation samples $(\mathbb{C}^1, \mathbb{C}^2, \dots, \mathbb{C}^M)$ of $M+W-1$ packets. In our method, we define the kernel density estimator as $\hat{\mathbf{f}}_\mathbf{l}$,

$$\hat{f}_l(\mathbb{C}) = \frac{1}{nh_l}\sum_{j=1}^{n} \mathrm{K}\left(\frac{\mathbb{C}-\mathbb{C}_{Nor}^{l,j}}{h_l}\right) \tag{4}$$

where K is the kernel function and h_l is the bandwidth. In particular, Epanechnikov quadratic kernel [14] is chosen owing to ensure the fairness of comparison to the counterpart RSS-based approach [2] and given by:

$$K(u) = \begin{cases} \frac{3}{4}(1-u^2), \text{if} |u| \leq 1 \\ 0, \text{otherwise} \end{cases} \tag{5}$$

h_l is a scaling factor that controls how wide the probability mass is spread around a point as well controls the smoothness or roughness of a density estimate. According to Scott's rule [15], the optimal bandwidth is given by:

$$h_l^* = 2.345\hat{\sigma}_l n^{-0.2} \tag{6}$$

where $\hat{\sigma}_l$ is an estimate for the standard deviation for $\mathbb{C}_{Nor}^{l,j}$.

For each RF link, we examine the cumulative distribution function (CDF) of the sample CSI correlation as $\hat{F}(\mathbb{C}_{Nor})$. It should be noted that $\mathbb{C}_{Abn}$ is always smaller than $\mathbb{C}_{Nor}$. Thus, the anomaly detection problem is equivalent to determine whether $\mathbb{C}$ is smaller than a lower bound $\hat{F}^{-1}(\beta)$ determined by a preset value β. The selection of β provides a tradeoff between false alarm and miss detection.

D. Position Estimation

In the previous two blocks, Pilot offlinely maintains a set of fingerprints into an abnormal radio map and continuously monitors the anomaly event in an online phase. As the target of locating the entities in real time, we now introduce the Position Estimation block.

Our objective is to accurately map the current fingerprint of the abnormal entity to the passive radio map during the online phase. The overall idea is to compare the obtained CSI measurements against the abnormal passive fingerprints database and thus selects the best match. Pilot chooses the maximum *a priori* probability (MAP) algorithm, which is a well-known probabilistic algorithm for performing fingerprint-based position estimation.

During the online localization stage, for an unknown location L where an abnormal entity is presented, we collect the abnormal CSI measurement $\mathbb{H}_{Abn}^l$ from link l. Let $\mathbb{L} = L_1, L_2, \cdots, L_m$ be the set of m locations on the passive radio map. Then, our position estimation task equates to find a location $L \in \mathbb{L}$ that maximizes *a priori* probability $P(\mathbb{H}_{Abn}^l|L)$. On the basis of Bayes' law, we formulate this optimization object function by:

$$L^* = \arg\max_L P(L|\mathbb{H}_{Abn}^l) = \arg\max_L [\frac{P(\mathbb{H}_{Abn}^l|L)P(L)}{P(\mathbb{H}_{Abn}^l)}] \tag{7}$$

Assume that all locations are equally probable, and $P(\mathbb{H}_{Abn}^l)$ is independent of location L, we have

$$L^* = \arg\max_L P(\mathbb{H}_{Abn}^l|L) \tag{8}$$

Since $\mathbb{H}_{Abn}^l$ is high-dimensional (52 subcarriers out of total 64 in 802.11n standard) and partially correlated, the statistical analysis of $\mathbb{H}_{Abn}^l$ is very complicated. Therefore, similar to anomaly detection, we consider the cross correlation of $\mathbb{H}_{Abn}^l$ over the fingerprints at each position L, and denote it as $\mathbb{C}_{Abn,L}^l$. We use kernel density to represent the probability in Eq.(9) whose calculation is similar to that in Eq.(4).

E. Data Fusion

If the detection points are deployed with high density, it is possible that a single motion can be detected by multiple links as shown in Figure 7. Therefore, we can extend our basic scheme to such scenarios in order to further improve the localization accuracy with data fusion method.

Given the measurements of multiple RF links, the final estimation is the position that the joint possibility is maximized, i.e.,

$$L^* = \arg\max_L \prod_l P(\mathbb{H}_{Abn}^l|L) \tag{9}$$

To reduce the computational complexity, only links with motion being detected will be included in the above calculation. Moreover, only the positions in the common area of the sensing zone of these links will be selected as the candidate for fingerprint mapping.

In summary, given the abnormal CSI measurement at each RF link, the kernel density-based MAP algorithm outputs the location L with maximal kernel density. Note that, we only consider the scenario when a single intruder exists, the localization of multiple intruders will be much more complicated and thus beyond the scope of this paper.

V. Performance Evaluation

In this section, we implement and evaluate Pilot in two typical indoor scenarios. We first describe our testbeds and data collection methodology in Section V-A. Afterwards, we validate the performance of anomaly detection and localization in Pilot, along with the comparison against state-of-art RSS-based approaches.

A. Implementation

I. Experimental Scenarios: We set up two typical indoor testbeds in Hong Kong University of Science and Technology as listed below:

1) ***Laboratory*** First, we performed experiments in a research laboratory covers an area of $7m \times 11m$. It is surrounded by various office facilities such as shelf, desk and chair, and therefore is subject to multipath effects. A total of 2 pairs of DPs and APs are placed according to the floor plan in Figure 8.
2) ***Lobby*** Second, we deployed Pilot a larger testbed in an "L"-shape lobby, which spreads over approximately $776m^2$. In this experimentation site, we symmetrically place two pairs of APs and DPs in opposite sides of parallelogram. The area of lobby is separated into a couple of squares and we choose 30 reference positions, each of them are $4m$ apart as shown in Figure 9.

II. Data Collection: In our experiments, we use TL-WR941ND router as wireless AP that transmit information over a RF link to DPs. Pilot DP is a standard HP laptop equipped with commercial 802.11n 5300 NICs, and the operating system is the Linux kernel 2.6.34. The Pilot server is running on one of these laptops. In our current implementation, we only use the first antenna and the enhancement with multiple antennas is left for our future work.

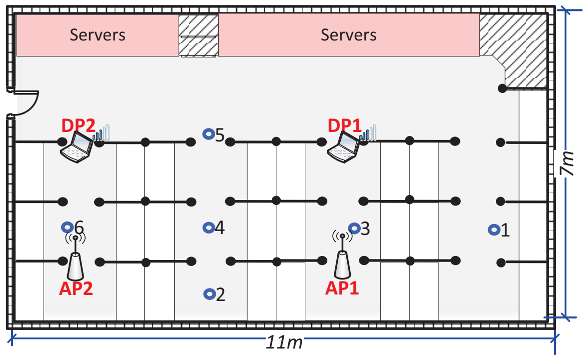


Fig. 8: Fingerprint Layout in Laboratory.

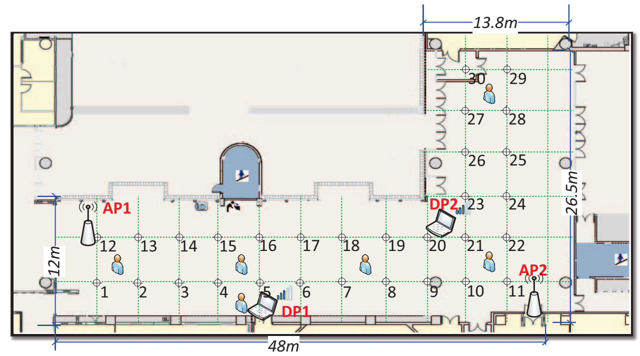


Fig. 9: Fingerprint Layout in Lobby.

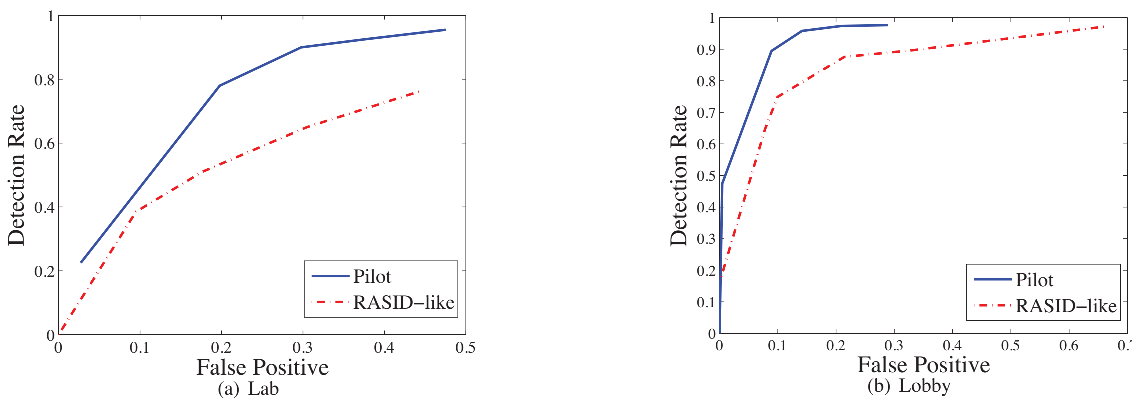


Fig. 10: Accuracy of Anomaly Detection.

During the whole localization period, APs will continuously send out beacon messages to DPs. DPs gather these messages along with CSIs and upload them to detection server for processing. In each scenario, we first collect measurement of every link in static environment without any person in the area, and construct normal profile. Then, we divide the geographic area of interest into uniform square grids, and mark some of them as crucial reference positions. Passive abnormal fingerprints are collected when one volunteer stand on each reference position. In this way, we construct passive radio maps in both lab and lobby.

B. Accuracy of Anomaly Detection

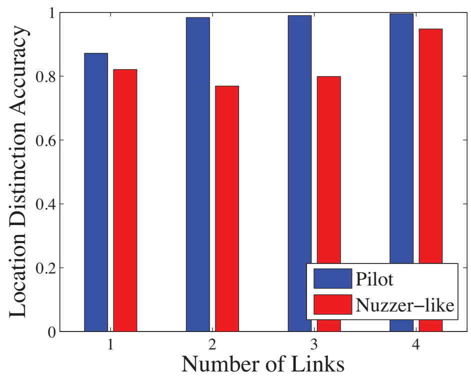
High accuracy of anomaly detection is necessary to guarantee the efficiency of device-free localization. In this section we evaluate whether Pilot can achieve this goal and conduct a comparison with the best RSS-based device-free motion detection system RASID [2].

Figure 10 plots the anomaly detection results in lab and lobby by using an Receiver Operating Characteristic (ROC) curve. This ROC curve graphically reveals the inherent tradeoff between the false positive (FP) rate and detection rate. FP rate (X-axis) also known as false alarm, represents the proportion that normal state is falsely detected as an anomaly. As mentioned in Section IV-C, the parameter β plays an important role in striking a balance of high detection rate with respect to low FP rate. In our experiment, we adapt β and fix the sliding window length to be 10. From the Lab scenario Figure 10(a), we have two obvious observations: 1) for a FP rate less than or equal to 10%, the detection rate of both approaches is alike to be around 40%; 2) for a FP rate greater than or equal to 30%, the detection rate of Pilot is about 90% that far exceeds RSS-based RASID. We have similar observations in the lobby scenario as shown in Figure 10(b). These results confirm that the CSI-based Pilot is superior to RSS-based approach in terms of anomaly detection.

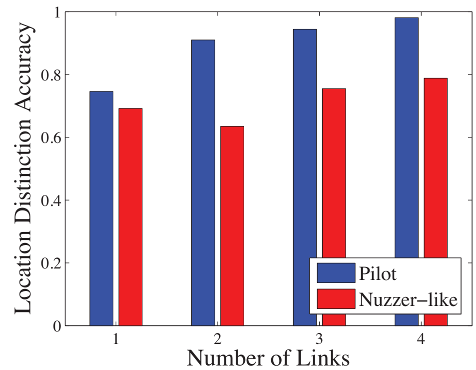
C. Accuracy of Localization

So far, we have described the performance of anomaly detection in a typical laboratory scenario. Clearly, a foremost criteria – localization accuracy is left for discussion in the following sections.

We analyze the location distinction accuracy of Pilot when there is only one single abnormal entity in Lab. In this scenari-

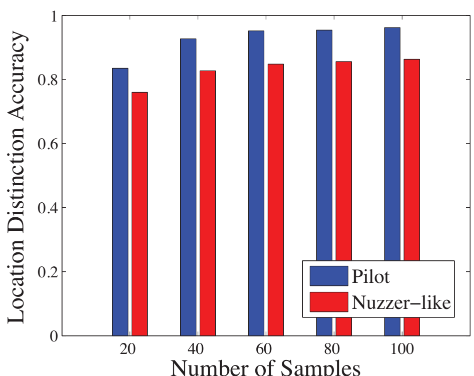


(a) Location Distinction in Lab

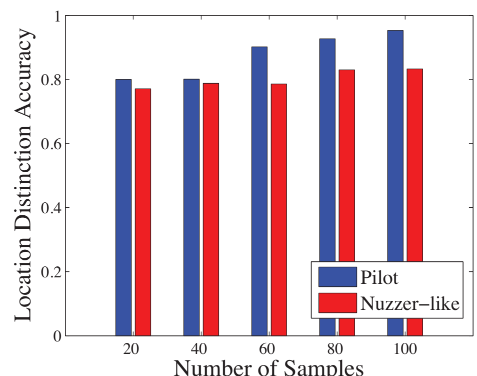


(b) Location Distinction in Lobby

Fig. 11: Single Entity Localization Accuracy with Different Link Numbers



(a) Location Distinction in Lab



(b) Location Distinction in Lobby

Fig. 12: Single Entity Localization Accuracy with Different Sample Numbers

o, Pilot deployment composes of 2 pairs of APs and DPs and 4 RF links in total. We evaluate the effect of increasing numbers of links (from 1 to 4) on the performance of the approach proposed in Pilot and the best RSS-based Nuzzer [3]. Instead of Gaussian distribution assumption, we use kernel density function to depict the distribution of RSS. Since kernel density is a better distribution approximation, the RSS-based system is Nuzzer-like but improved one. These results are presented in Figure 11(a), such that if only 1 link is available in the area, Pilot achieves higher accuracy (6%) than Nuzzer-like approach. As changing the links from 2 to 4, the accuracy of Pilot can extend to over 90% even to 98% while Nuzzer-like approach only has a slight improvement. In other words, the use of more RF links would lead to an obvious increase in location distinction accuracy of CSI-based Pilot rather than RSS-based Nuzzer. This indicates the benefits of Pilot over RSS-based method in gaining distinction capability. We also study the location distinction performance of a single abnormal entity in Lobby. Similar performance is achieved with similar deployment as shown in Figure 11(b). Comparing Figure 11(a) and Figure 11(b), we can observe that our system performs better in Lab scenario due to the more abundant multipath reflections.

Figure 12(a) is a result in Lab that compares the precision of Pilot and Nuzzer with respect to different numbers of samples collected as fingerprints. This figure shows that increasing sample number will bring in higher accuracy in some extent, because with more fingerprints, we can better approximate the distribution of CSI correlation, and results in lower fingerprint mapping error according to (8). More than 10% percents accuracy improvement can be obtained with more samples in our experiment. However, more samples means longer offline phase is required for map construction. Therefore, there is a tradeoff between the data collection time and localization accuracy. It is also shown that the proposed CSI-based Pilot always outperforms the corresponding RSS-based Nuzzer with respect to localization accuracy. Similar performance is achieved in the Lobby scenario as shown in Figure 12(b).

From empirical experiments in these two scenarios, we can conclude that frequency diversity of CSI helps Pilot outperform the RSS-based scheme and such advantage is obvious when more RF links are available.

VI. Conclusions and Future Work

In this paper, we present Pilot, a CSI-based passive device-free indoor fingerprinting system in WLAN. It is the first proposal to leverage temporal stability and frequency diversity characteristics of CSI for developing a "passive" fingerprint for device-free localization. In contrast to the traditional fingerprint approach, we integrate an Anomaly Detection block to facilitate the device-free nature. We apply the kernel density-based approach to calculate the CSI correlation and thus detect abnormal entities. Moreover, we develop another block for estimating the location of target, and analysis the feasibility of distinguish multiple entities simultaneously. We implement Pilot with commercial IEEE 802.11n NICs and evaluated its performance in two different indoor scenarios. Pilot outperforms the RSS-based RASID system on anomaly detection and Nuzzer system on localization with the same infrastructure deployment. The high accuracy of Pilot system demonstrates its potential to dramatically improve the performance of existing location-dependent applications.

In the perspective of future research, we intend to exploit other approaches other than fingerprinting for CSI-based device-free indoor localization to offload the calibration efforts, and improve accuracy as well. Moreover, multiple antennas can also be leveraged to achieve better system performance.

Acknowledgment

This research is supported in part by Guangdong Natural Science Funds for Distinguished Young Scholar (No. S20120011468), Hong Kong RGC Grant HKUST617212, Guangzhou Pearl River New Star Technology Training Project (No.2012J2200081), Guangdong NSF Grant (No.S2012010010427), China NSFC Grant 61202454.

References

[1] M. Youssef, M. Mah, A. Agrawala,"Challenges: Device-free Passive Localization for Wireless Environments," in *Proc. of ACM MobiCom*, 2007.

[2] A.E. Kosba, A. Saeed, M. Youssef,"RASID: A Robust WLAN Device-free Passive Motion Detection System," in *Proc. of IEEE PerCom*, 2012..

[3] M. Seifeldin, A. Saeed, A.E. Kosha, A. EI-Keyi and M. Youssef,"Nuzzer A Large-Scale Device-Free Passive Localization System for Wireless Environments," in *Proc. of IEEE Transactions on Mobile Computing*, 2011.

[4] N. Patwari,"Robust Location Distinction using Temporal Link Signatures," in *Proc. of ACM MobiCom*, 2007.

[5] J. Zhang, M.H. Firooz, N. Patwari, S.K. Kasera,"Advancing Wireless Link Signatures for Location Distinction," in *Proc. of ACM MobiCom*, 2008.

[6] J. Zhang, S.K. Kasera, N. Patwari, P. Rai,"Distinguishing Locations Across Perimeters Using Wireless Link Measurements," in *Proc. of IEEE INFOCOM*, 2011.

[7] Z.G. Zhong, D. Wang, and T. He, "Sensor Node Localization Using Uncontrolled Events," in *Proc. of IEEE ICDCS*, 2008.

[8] D. Zhang, J. Ma, Q. Chen, and L.M. Ni,"An RF-Based System for Tracking Transceiver-Free Objects," in *Proc. of IEEE PerCom*, 2007.

[9] D. Zhang, Y.H. Liu, L.M. Ni,"RASS: A real-time, accurate and scalable system for tracking transceiver-free objects," in *Proc. of IEEE PerCom*, 2011.

[10] X.N. Guo, D. Zhang, Lionel M. Ni, "Localizing Multiple Objects in an RF-based Dynamic Environment," in *Proc. of IEEE ICDCS*, 2012.

[11] A. Bhartia, Y. Chen, S. Rallapalli, and L. Qiu, "Harnessing Frequency Diversity in Wi-Fi Networks," in *Proc. of ACM MobiCom*, 2011.

[12] J. Yang, Y. Ge, H. Xiong, Y.Y. Chen, and H.B. Liu, "Performing joint learning for passive intrusion detection in pervasive wireless environments," in *Proc. of IEEE INFOCOM*, 2010.

[13] A. R. Ansari and R. A. Bradley, "Rank-sum test for dispersions," in *Proc. of Ann. Math. Statist*, 1174-1189, 1960.

[14] Epanechnikov, V.A., "Non-parametric estimation of a multivariate probability density," in *Proc. of Theory of Probability and its Applications*, 1969.

[15] D. W. Scott, "Multivariate Density Estimation: Theory, Practice, and Visualization,", in *Proc. of John Wiley & Sons, Inc.*, 1992.

[16] Want, R. and Hopper, A. and Falcão, V. and Gibbons, J., "The active badge location system," in *Proc. of ACM Transactions on Informaton Systems*, 1992.

[17] Y. Chen, D. Lymberopoulos, J. Liu, B. Priyantha, "FM-based Indoor Localization," in *Proc. of ACM MobiSys*, 2012.

[18] J. Krumm, L. Williams, and G. Smith, "SmartMoveX on a Graph - An Inexpensive Active Badge Tracker," in *Proc. of UbiComp*, 2002.

[19] Ni, L. M. and Liu, Y. and Lau, Y. C. and Patil, A.,"LANDMARC: indoor location sensing using active RFID," in *Proc. of IEEE PerCom*, 2003.

[20] Youssef, M. and Agrawala, A., "The Horus WLAN location determination system," in *Proc. of ACM MobiSys*, 2005.

[21] D. Zhang, L.M. Ni, "Dynamic Clustering for Tracking Multiple Transceiver-free Objects," in *Proc. of ACM SenSys*, 2009.

[22] Y. Zhao and N. Patwari, "Noise Reduction for Variance-Based Device-Free Localization and Tracking," in *Proc. of IEEE SECON*, 2007.

[23] C. Wu, Z. Yang, Y.H. Liu, and W. Xi, "WILL: Wireless Indoor Localization without Site Survey," in *Proc. of IEEE INFOCOM*, 2012.

[24] K. Wu, J. Xiao, Y. Yi, and Lionel M. Ni, "FILA: Fine-grained Indoor Localization," in *Proc. of IEEE INFOCOM*, 2012.

[25] J. Xiao, K. Wu, Y. Yi, and Lionel M. Ni, "FIMD: Fine-grained Device-free Motion Detection," in *Proc. of IEEE ICPADS*, 2012.

[26] J. Wilson and N. Patwari, "See-Through Walls: Motion Tracking Using Variance-Based Radio Tomography Networks," in *Proc. of IEEE Transactions on Mobile Computing*, 2011.

[27] Priyantha, N. B. and Chakraborty, A. and Balakrishnan, H., "The Cricket location-support system," in *Proc. of MobiCom*, 2000.

[28] P. Bahl and V. N. Padmanabhan, "RADAR: An inbuilding RF-based user location and tracking system," in *Proc. of IEEE INFOCOM*, 2000.

[29] J. Wilson and N. Patwari, "Radio Tomographic Imaging with Wireless Networks," in *Proc. of IEEE Transactions on Mobile Computing*, 2010.

[30] Halperin, D. and Hu, W. J. and Sheth, A. and Wetherall, D., "Predictable 802.11 packet delivery from wireless channel measurements," in *Proc. of ACM SIGCOMM*, 2010.

[31] R. S. Moore, R. Howard, P. Kuksa, and R. P.Martin, "A Geometric Approach to Device-Free Motion Localization Using Signal Strength," in *Technical Report, Rutgers University*, 2010.