Wireless Netw
DOI 10.1007/s11276-008-0140-3

Zero-configuration indoor localization over IEEE 802.11 wireless infrastructure

Hyuk Lim · Lu-Chuan Kung · Jennifer C. Hou · Haiyun Luo

Abstract With the technical advances in ubiquitous computing and wireless networking, there has been an increasing need to capture the context information (such as the location) and to figure it into applications. In this paper, we establish the theoretical base and develop a localization algorithm for building a zero-configuration and robust indoor localization and tracking system to support location-based network services and management. The localization algorithm takes as input the on-line measurements of received signal strengths (RSSs) between 802.11 APs and between a client and its neighboring APs, and estimates the location of the client. The on-line RSS measurements among 802.11 APs are used to capture (in real-time) the effects of RF multi-path fading, temperature and humidity variations, opening and closing of doors, furniture relocation, and human mobility on the RSS measurements, and to create, based on the truncated singular value decomposition (SVD) technique, a mapping between the RSS measure and the actual geographical distance. The proposed system requires zero-configuration because the on-line calibration of the effect of wireless physical characteristics on RSS measurement is automated and no on-site survey or initial training is required to bootstrap the system. It is also quite responsive to environmental dynamics, as the impacts of physical characteristics changes have been explicitly figured in the mapping between the RSS measures and the actual geographical distances. We have implemented the proposed system with inexpensive off-the-shelf Wi-Fi hardware and sensory functions of IEEE 802.11, and carried out a detailed empirical study in our departmental building, Siebel Center for Computer Science. The empirical results show the proposed system is quite robust and gives accurate localization results.

Keywords Location estimation · Singular value decomposition · Signal strength based localization · IEEE 802.11 wireless infrastructure

A preliminary version of this paper appeared in IEEE INFOCOM 2006, Barcelona, Spain, April 2006 [1].

H. Lim (✉)
Department of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500-712, South Korea
e-mail: hlim@gist.ac.kr

L.-C. Kung · J. C. Hou · H. Luo
Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana-Champaign, IL 61801, USA
e-mail: kung@cs.uiuc.edu

J. C. Hou
e-mail: jhou@cs.uiuc.edu

H. Luo
e-mail: haiyun@cs.uiuc.edu

1 Introduction

We have witnessed swift advances in wireless communication and networking over the last decade. Such advances have made it possible to realize the notion of ubiquitous computing and communications. As a result, the capability for capturing contexts and figuring them into the computing/communication process has become an immediate need. As the physical location is one of the most important context parameters, its availability, as predicated independently in several market surveys [2, 3], will define the emerging market of location-based services with $7B to $8B projected revenue over the next few years.

Availability of physical locations of wireless-enabled computing devices also facilitates wireless resource management. The proliferation of wireless-enabled devices (and applications built on top of them) has aggravated the competition for limited bandwidth in the underlying wireless infrastructure. This is particularly true for devices operating in the unlicensed frequency band such as IEEE 802.11-enabled devices [4]. Without adequate resource provisioning, wireless-enabled devices may soon become the victims of their own success. A thorough understanding of how wireless resources are spatially and temporally used is thus necessary, based on which intelligent management and provisioning policies can be developed. As the limited frequency band for wireless communication is essentially a spatial resource defined *by the location*, the availability of device location information will enable advanced management and control of wireless networks.

Building an indoor localization system to estimate the location of wireless devices is a challenging research problem. In an indoor environment, RF signal propagation is affected by a number of factors such as multi-path fading, temperature and humidity variations, opening and closing of doors, furniture relocations, and the presence and mobility of human beings. To demonstrate the impact of these factors, we placed a stationary 802.11 monitor in one of the student cubicles and logged the received single strength (RSS) from the nearest 802.11b access point during the period of January 1–7, 2005. As shown in Fig. 1(a), the RSS, averaged over a time window of 5 min, varied at both small (seconds) and large (hours) time granularities, and the variation was as high as 7 dBm. Figure 1(b) gives a zoom-in view of the RSS measurement on January 4. From these two figures we observe substantially higher RSS variations during weekdays (January 3–7) and during office hours (9:00 am–6:00 pm). The high variation of RSSs excludes localization algorithms that require a full-scale, on-site survey and training process to initialize the localization system (e.g., [5–9]). These algorithms do not scale to large environments such as enterprise buildings and factory floors because they require extremely cumbersome and intrusive re-calibrations in order to maintain high accuracy in the presence of environment changes (as discussed above). A localization system should adapt to all environmental dynamics, and be able to auto-configure itself in response to such environmental dynamics. It should also be resilient to measurement errors (introduced in the sensory inputs from, for example, off-the-shelf 802.11 interfaces).

Several R&D efforts have been made in utilizing the physical location for detecting rogue APs and for gross-grained localization [5, 7, 8, 10–13]. While these efforts have demonstrated certain level of feasibility and inspired the work proposed in this paper, there are still several research issues that should be addressed, before a low-cost, adaptive, and robust localization and tracking system can eventually be realized. In particular, existing approaches either rely on special sensory hardware [10–12, 14], do not achieve fine localization granularity [6–8, 13], or require extensive a priori configuration and training [6–9]. (We will provide a more detailed summary of existing work in Sect. 2.) One of the most important issues is the lack of the on-line, automated calibration mechanism that captures the relationship between the RSS measurements and the geographical distances in both the time and spatial domains and the theoretical base that provides, based on the captured relationship, an accurate mapping between the RSS measurements and the geographical distances.

In this paper, we attempt to build a *zero-configuration, robust, indoor localization and tracking* system to support location-based network services and management. Given the phenomenal popularity of indoor 802.11 (a.k.a Wi-Fi) WLAN deployments and 802.11-enabled devices, we leverage the Wi-Fi infrastructure and augment its functionality and application domain from wireless connectivity to fine-grained localization. Our system is grounded on the innovation that turns the off-the-shelf 802.11-enabled devices into sensors, the 802.11 access points into anchor nodes, and the managed 802.11 infrastructure into a wireless sensor network.

Specifically, we have developed a localization algorithm that takes as input the on-line measurements of RSSs (i) between 802.11 APs and (ii) between a client and its

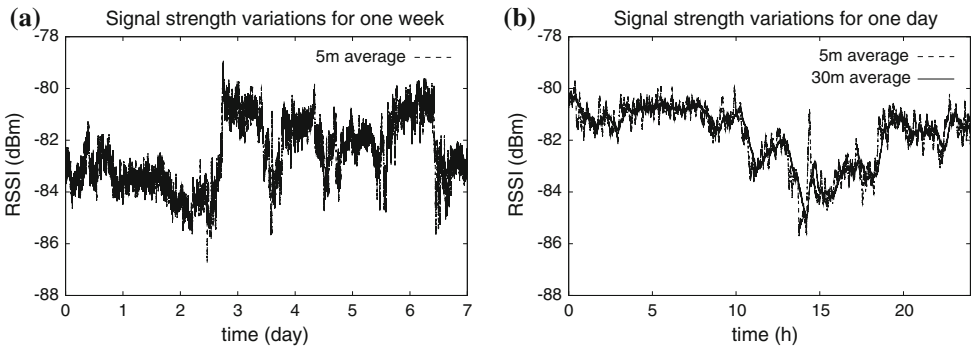


Fig. 1 Received signal strength (RSS) measurements that show large variations over time at different time scale. (**a**) Five-minute average RSS, January 1–7, 2005, (**b**) 5/30-min average RSS, January 4, 2005

neighboring APs, and estimates the location of the client. The on-line RSS measurements between 802.11 APs are used to capture (in real-time) the effects of environmental dynamics (e.g., RF multi-path fading, temperature and humidity variations, and human mobility) on the RSSs, and to create a mapping between the RSS measure and the actual geographical distance. The mapping is created with *truncated singular value decomposition* (SVD) techniques, with the objectives of mitigating the effect of measurement errors, while retaining as much environmental information as possible. The location of a wireless client can be calculated with the signal-distance mapping (SDM) and the user-measured RSSs between itself and its neighboring APs.[1]

The proposed system possesses several desirable features. First, it does not require any additional proprietary hardware. Second, it requires zero-configuration because the on-line calibration is automated by on-line measurements between APs. Third, it is resilient to measurement noises due to the truncated SVD technique used for computing the signal-distance mapping. We have carried out an empirical study in our departmental building, Siebel Center for Computer Science, with the use of Linksys WRT54G wireless routers equipped with customized Kismet 802.11 layer-2 sniffer software. The empirical results indicate the proposed system is quite robust and responsive to environmental dynamics, and gives accurate localization results (i.e., with the localization error within 3 m). This is not surprising, as the impacts of environmental dynamics have been *explicitly* figured in the mapping between the RSS measures and the actual geographical distances.

It has come to our attention recently that Gwon and Jain [15] has shared the same view as we do, and proposed a calibration-free scheme, called *Proximity in Signal Space* (PSS). In PSS, each AP collects RSSs from all the other APs and generates multiple linear functions, each representing a RSS vs. distance mapping between itself and the corresponding AP. In spite of its simplicity, the method which PSS used to generate the mapping is heuristic-based, and is more susceptible to the location of a wireless client to be localized. As a result, the localization accuracy degrades if a client is not close enough to one of APs. As reported in the empirical study (Sect. 4), the localization error incurred in PSS is almost twice of that incurred in SDM in most cases.

The rest of the paper is organized as follows. In Sect. 2, we present a summary of existing, related work, highlight the limitation of the current state-of-art approaches, and motivate the need to construct a better signal-distance mapping. In Sect. 3, we first give the system architecture, and then delve into the theoretical base and the implementation details. This is then followed by a discussion on several deployment issues, such as where to place wireless monitors, how to deal with co-existence of IEEE 802.11a/b/g, and the trade-off between client privacy and infrastructure security. In Sect. 4, we elaborate on how we customize the Kismet 802.11 layer-2 sniffer software and carry out the empirical study. We also report our empirical findings there. Finally, we conclude in Sect. 5 the paper with several potential avenues for future work.

[1] The current 802.11 client interface has already made such measurements to locate and associate with the AP with the strongest beacon signal.

2 Related work

In this section, we give a succinct summary of existing work and discuss the limitation of the current state of art. We categorize existing work roughly into two groups: *indoor localization* and *localization with the use of wireless sensor networks*.

2.1 Indoor localization

Systems designed for indoor localization can be classified based on the signal types (infrared, ultrasound, ultra-wideband, and radio frequency), signal metrics (AOA: angle of arrival, TOA: time of arrival, TDOA: time difference of arrival, and RSS: received signal strength), and the metric processing methods (triangulation and scene profiling).

Active Badge [16] is one of the early systems based on infrared ranging. As the infrared signal requires a line-of-sight propagation path between the transmitter and the receiver, objects in-between can easily block the signal and deteriorate the performance of the system. Its successor Active Bat [17] and MIT Cricket [12], [18] are based on the ultrasound technology. The major problem with the ultrasound technology is that the propagation velocity of the ultrasound is easily affected by the temperature and humidity, which introduce ranging errors over the long term. Ubisense localization system [11] represents a recent advance in leveraging ultra-wideband (UWB) for localization. It achieves fine-grained indoor localization with the accuracy of up to $\sim$6 in (15 cm). However, the cost for a Ubisense UWB reader is currently more than an order of magnitude higher than that of an 802.11 AP.

Widespread deployment of the 802.11 infrastructure and the ubiquity of Wi-Fi embedded devices are the most compelling reason for radio-frequency based indoor localization. Systems based on AOA [14], TOA [10] or TDOA [12, 19] have been proposed and have reportedly achieved up to 1-m localization granularity. However, the measurements of AOA, TOA, or TDOA necessitate special

hardware at either the infrastructure side or the client side. For example, the AOA measurement requires the use of directional antenna with beam forming, while TOA and TDOA measurements require the availability of fine-grained timers (10 ns corresponding to 3-m RF localization precision).

Since received signal strength (RSS) measurement is based on a sensory function already available in most 802.11 interfaces, RSS-based indoor localization therefore receives significant attention from both academia and industry. Existing solutions can be further categorized according to their signal processing methods. Range-based approaches collect RSS measurements, estimate the distances between a client and reference points (usually 802.11 APs), and then apply the triangulation method to derive the client location [13, 20, 21]. Madigan et al. [22] build a Bayesian hierarchical model to make a tradeoff between the training dataset sizes and the levels of the range estimation precision. Approaches in the other category establish a location-RSS mapping through scene profiling, and infer the client's coordinates using different matching algorithms [6–8, 23–25].

While these RSS-based efforts have inspired the proposed work, there is still room for performance improvement, especially with respect to the adverse impact of RSS dynamics. Recall that in Sect. 1, we observe a wide variation (up to 20 dBm) of RSS measurements in typical 802.11 infrastructure, and the trend persists at different time scales. As these dynamics are caused by a number of environmental factors, and asymmetric non-Gaussian noise [23, 26–28], *the assumption that the indoor space remains consistent from the training phase to the localization phase thus does not hold true in practice*. Supervised re-initialization could be extremely cumbersome and intrusive, and does not scale to large enterprise buildings and factory floors. On-line training [7] requires cooperation of autonomous users, and may not be a desirable solution. What seems to be reasonable is a system that *acknowledges* the existence of RSS dynamics and incorporates a *fully-automated*, on-line calibration mechanism to characterize (but not model in close forms) their relation with environmental factors in both the time and spatial domains.

Working toward this end, Gwon and Jain [15] proposed a calibration-free technique, called *Proximity in Signal Space* (PSS), that uses inter-AP RSS measurements. Each AP collects RSSs from all other APs and generates multiple linear functions, each representing a RSS vs. distance mapping between itself and the corresponding AP. For example, in an environment with 4 APs, each AP establishes three linear functions. A client node then uses (i) the mapping linear functions kept at the closest AP (i.e., the AP with the strongest RSS) to compute its distances to all the APs except the closest AP, and (ii) the mapping function kept at the second closest AP to compute its distance to the closet AP. In this fashion, PSS gives an accurate approximation when a client is close enough to the closest AP. The authors also proposed a lateration algorithm called *Triangular Interpolation and eXtrapolation (TIX)*. The calibration-free TIX algorithm was reported to achieve mean distance error within 5.4 m.

2.2 Wireless sensor network localization

Localization techniques that leverage sensor networks can be classified into either range-based or range-free approaches. Range-based approaches measure the proximity (in terms of hop-count or estimated distance) to a few landmarks (or beacons) with known locations, and apply different algorithms/strategies to infer sensor node coordinates [29–36]. These proposals work well in the *isotropic* environments (where the proximity measure, such as the hop count, well approximates the Euclidean distance in *all* directions), but their performance severely degrades in *anisotropic* environments.

Range-free approaches use wireless reachability that simply indicates whether or not a node is located within a radio range of other node and estimate the position of a client node without computing the geographic distances to beacon nodes [37–39]. As these methods do not estimate the geographic distances between nodes, they are less sensitive to measurement noises, and a lateration task is not necessary. However, the performance may also be marginal in indoor environments because a high node density is required to achieve accurate location estimation.

3 Indoor localization system

3.1 System overview

The proposed system is composed of 802.11 access points and off-the-shelf 802.11-compliant devices. As shown in Fig. 2, the system architecture follows a typical 802.11 WLAN deployment in indoor environments. At the infrastructure side, APs with known locations serve as *anchors*. These APs record the received signal strength (RSS) of beacon broadcasts from each other. In addition, inexpensive wireless monitors (augmented with the customized Kismet software [40]) can be deployed in under-provisioned areas to serve as additional anchor points and provide additional measurements. Choices on the number of additional wireless monitors and where to place them create extra control knobs for improving localization accuracy. We will elaborate on this issue in Sect. 3.4. The inter-AP (or inter-anchor node) RSS measurements are made periodically to realize fully automated and on-line

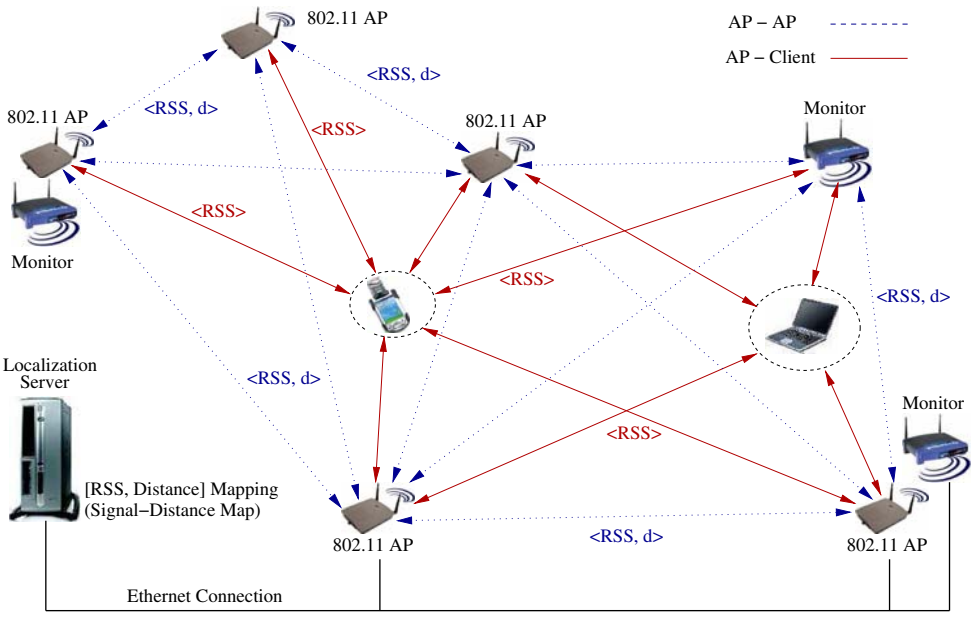


Fig. 2 System architecture for indoor localization

calibration of RSSs in the spatio-temporal domain. A mapping that characterizes the relationship of RSS measurements and geographical distances to anchors is then created *on-line* with an SVD-based approach. Depending on the mode in which the system operates, either a client or the infrastructure measures the RSSs between the client and its neighboring anchor nodes, applies the signal-distance mapping, infers the client's geographical distances to anchor nodes, and estimates its location via triangulation/trilateration.

3.2 Zero-configuration robust localization

As the RSS measure is susceptible to several physical wireless characteristics, such as multi-path fading, and signal attenuation due to the changes in temperature, humidity, and object mobility, it does not correlate in a simple and stable fashion to the geographic distance. In fact, the correlation structure changes in the spatial-temporal domains, and cannot always be captured in nice and close-form theoretical models. Nevertheless, their accurate characterization is crucial to accurate localization. We propose a new, theoretically-grounded technique to analyze the relationship between the geographic distance and the RSS measure.

3.2.1 Localization problem

Conceptually, localization can be considered as an embedding problem that maps the set of objects into an embedding space. In Lipschitz embeddings, a coordinate space is so defined that it corresponds to a *reference set* of objects (i.e., the anchor nodes consisting of APs and wireless monitors in the vicinity of a client to be localized), and the coordinate values of an object o are the distances from o to the reference objects [41–43]. For example, if an object is apart from two reference objects by 1 and 2 (units of distance), respectively, it has the coordinate of $[1, 2]^T$ in the embedding space. Based on this concept, each client to be localized has two coordinates in Lipschitz embedding spaces that correspond, respectively, to the RSS measure and the Euclidean distance between itself and anchor nodes.

Consider a Wi-Fi infrastructure in which m anchor nodes (i.e., APs and/or wireless monitors) exist in the vicinity of a client node (whose position is yet to be localized). The locations of anchor nodes and the client node n are denoted as $\mathbf{x}_i \in \mathbb{R}^d$ for $i = \{1, \ldots, m\}$ and $\mathbf{x}_n \in \mathbb{R}^d$ in a d-dimensional space, respectively. The geographic distance between two nodes, $\mathbf{x}_i$ and $\mathbf{x}_j$ is then defined by the Euclidean distance:

$$d_{ij} = f_d(\mathbf{x}_i, \mathbf{x}_j) := \sqrt{\sum_{k=1}^{d} (x_{ik} - x_{jk})^2}, \tag{1}$$

where x_{ik} and x_{jk} are the kth coordinates of $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively. Let s_{ij} be the RSS measured by the ith node to the jth node. Then the localization problem we consider can be formally stated as:

$$\begin{aligned} &\textit{Given} : \mathbf{x}_i, s_{ij}, \text{and } s_{ni} \text{ for } i, j \in \{1, \ldots, m\}, \\ &\quad \textit{Estimate} : \mathbf{x}_n \text{ for a client node } n. \end{aligned}$$

That is, under the assumption that locations $\mathbf{x}_i$ of the anchor nodes are known, the problem is to estimate, with the use of the RSS measurements s_{ij} and s_{ni} for $i, j \in \{1, \ldots, m\}$, the geographic location $\mathbf{x}_n$ of client n.

3.2.2 Anisotropic indoor environment

Let $f_p : \mathbb{R}^{2d} \longmapsto \mathbb{R}$ denote the function that describes the mapping from the geographic locations ($\mathbf{x}_i$ and $\mathbf{x}_j$) to the measured RSS s_{ij} for each pair of nodes, where the RSS is written as $s_{ij} = f_p(\mathbf{x}_i, \mathbf{x}_j)$. The environment is said to be *isotropic* if the mapping $f_p(\mathbf{x}_i, \mathbf{x}_j)$ is a function of the Euclidean distance between $\mathbf{x}_i$ and $\mathbf{x}_j$, i.e., $s_{ij} = f_p(\mathbf{x}_i, \mathbf{x}_j) = g_p(d_{ij})$ and $g_p : \mathbb{R} \longmapsto \mathbb{R}$.

In practice, the RSS measured by a node is subject to the physical wireless characteristics in indoor environments and differs in different directions and locations. This implies that the RSS between a pair of nodes depends greatly on the positions at which these nodes are located, and the environment is *anisotropic*. For example, two wireless devices that are apart from an AP by the same distance can have quite different RSS measurements due to different intermediate obstacles between the devices and the AP.

In order to obtain accurate localization results, one may record RSSs at every possible location through off-line training process and compare the RSS of a client node to be localized with the recorded RSSs to find the best matching RSS pattern. As mentioned in Sect. 1, this approach is not desirable. Instead, we characterize the anisotropic relationship between the geographic distance and the RSS automatically and in real time.

3.2.3 Characterizing the relation between the RSS and the geographic distance

We now analyze the RSSs measured between anchor nodes with known geographic locations, and derive the *signal-distance mapping* (SDM) that describes the relation between the RSSs and the geographic distances. Specifically, the RSSs measured from a node to anchor nodes define the coordinate of the node. Given that there exist m anchor nodes, the coordinate of a node i in an m-dimensional Lipschitz embedding space is represented by the signal vector:

$$\mathbf{s}_i = [s_{i1}, \ldots, s_{im}]^T$$

where s_{ij} is the strength of the signal emitted by the jth anchor node and received by node i. In our implementation, RSS values are measured in dBm and $s_{ij} = -\mathrm{RSS}_{ij}$. The overall embedding space for RSSs measured between anchor nodes can be represented by an m-by-m RSS matrix $\mathbf{S}$, whose ith column is the coordinate of the ith beacon node:

$$\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_m]. \quad (2)$$

Similarly, we define the geographic distance vector and matrix as $\mathbf{d}_i = [d_{i1}, \ldots, d_{im}]^T$, and $\mathbf{D} = [\mathbf{d}_1, \ldots, \mathbf{d}_m]$. The geographic matrix $\mathbf{D}$ is an m-by-m symmetric square matrix with zero diagonal entries.

Example Consider an indoor localization system consisting of four APs (monitors), of which the locations are (0,0), (5,0), (0,5), and (5,5), respectively. Then, the geographic matrix $\mathbf{D}$ is given by

$$\mathbf{D} = \begin{bmatrix} 0 & 5 & 5 & 5\sqrt{2} \\ 5 & 0 & 5\sqrt{2} & 5 \\ 5 & 5\sqrt{2} & 0 & 5 \\ 5\sqrt{2} & 5 & 5 & 0 \end{bmatrix}.$$

Also assume (for demonstration purpose) that the transmit power of AP is 20 mW (13 dBm) and the path loss in the indoor environment follows the log-distance path loss model as follows:

$$PL = PL(d_0) + 10\alpha \log\left(\frac{d}{d_0}\right),$$

where $PL(d_0)$ is the path loss at a close-in reference distance d_0, and α is the path loss exponent. We assume there is 30 dB attenuation at $d_0 = 1$ m and $\alpha = 2.6$ at an office with hard partitions in [44]. Then, the RSS matrix $\mathbf{S}$ is given by

$$\mathbf{S} = \begin{bmatrix} s_{11} & 35.2 & 35.2 & 39.1 \\ 35.2 & s_{22} & 39.1 & 35.2 \\ 35.2 & 39.1 & s_{33} & 35.2 \\ 39.1 & 35.2 & 35.2 & s_{44} \end{bmatrix},$$

where s_{ii} is so-called *self-RSS*, which is measured between two co-located anchors. (We will further elaborate on this in Sect. 3.3) Note that $PL(5) = 48.2$ dB and $PL(5\sqrt{2}) = 52.1$ dB.

A good candidate for the signal-to-distance mapping (SDM) is the optimal linear transformation $\mathbf{T}$, where the geographic distance from a client node to an anchor node is represented as a weighted sum of the RSSs to *all* the anchor nodes. The weights are the elements of $\mathbf{T}$, which is the *least square solution* that minimizes the discrepancy between the geographic distance and the computed distance. An important issue for this transformation is whether there exists a linear relationship between the RSS (measured in dBm) and the geographic distance. To this end, we use a *linearizing function* $\mathcal{F}_d(d) = \log(d)$, with the rationale that indoor path loss obeys the distance–power law [44].[2] Note that since RSSs are already measured in dBm, the SDM can be expressed on logarithmic scale by taking the logarithmic function on the RSSs. On the other hand, the propagation models addressed above do not hold for small values of the distance. If the distance to be estimated is small, it is possible to use the first-order approximation $\mathcal{F}_d(d) = d$ as

[2] Our algorithm does not rely on a specific propagation model for indoor environments. Instead, we exploit the characteristics that the signal strength is inversely proportional to the distance to the power of a path loss exponent for constructing the SDM.

the linearizing function. This case corresponds to an exponential attenuation of RF signals with respect to the distance because the RSSs is measured in dBm. (We will compare the performance between the two linearizing functions in Sect. 4.)

The transformation matrix $\mathbf{T}$ is an m-by-m square matrix, of which the element t_{ij} represents the effect of the RSS measurement to the jth anchor node on the geographic distance to the ith anchor node, and can be considered as a scaling factor (weight) from the RSS to the geographic distance. Each row vector $\mathbf{t}_i$, which corresponds to the weight for computing the distance to the ith anchor node, can be obtained by minimizing the following objective function:

$$e_i = \sum_{k=1}^{m} (\mathcal{F}_d(d_{ik}) - \mathbf{t}_i \mathbf{s}_k)^2 = ||\mathcal{F}_d(\mathbf{d}_i^T) - \mathbf{t}_i \mathbf{S}||^2. \quad (3)$$

If the logarithm function is used as a linearizing function in the above equation, the value of the distance from oneself (so-called *self-distance*) is $-\infty$ (i.e., $d_{ii} = 0$) and will be replaced with a small value. (We will further elaborate on this in Sect. 3.3.) After some algebraic operations, we can readily obtain that the least-square solution for the row vector $\mathbf{t}_i$ is

$$\mathbf{t}_i = \mathcal{F}_d(\mathbf{d}_i^T)\mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1}.$$

That is, we can express the SDM as

$$\mathbf{T} = \mathcal{F}_d(\mathbf{D})\mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1}. \quad (4)$$

Note that SDM retains the effects of physical wireless characteristics on the RSS to *all* the anchor nodes, and hence well characterizes the anisotropic relationship between RSSs and geographic distances.

Now a client n with an unknown position measures the RSSs between itself and its neighboring anchor nodes (i.e., $\mathbf{s}_n$), and computes the geographic distance from the client to neighboring anchor nodes by matrix multiplication, (i.e., $\mathbf{d}_n = \mathcal{F}_d^{-1}(\mathbf{T}\mathbf{s}_n)$). The geographic location of the client n can be determined by lateration algorithms [29–32].

3.2.4 Improving robustness of SDM to measurement noises

The RSSs are susceptible to measurement noises introduced in the sensory inputs from, for example, off-the-shelf 802.11 interfaces. These noises may become manifest in the computation of the matrix inversion in Eq. 4. To reduce such adverse effects, we use the *truncated SVD pseudo-inverse* method described in [45]. Let the SVD of $\mathbf{S}$ be expressed as

$$\mathbf{S} = \mathbf{U} \cdot \begin{bmatrix} \mathbf{\Sigma} & 0 \\ 0 & 0 \end{bmatrix} \cdot \mathbf{V}^T,$$

where $\mathbf{\Sigma}$ is a diagonal matrix: $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_W)$, $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_m]$ and $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$ are column and row orthogonal matrices, and the subscript W is the rank of matrix $\mathbf{S}$ [46]. The σ_i's are singular values of $\mathbf{S}$ in the decreasing order (i.e., $\sigma_1 \geq \cdots \geq \sigma_W > 0$). The pseudo-inverse of the matrix $\mathbf{S}$, $\mathbf{S}^+$, is written as

$$\mathbf{S}^+ = \mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-1} = \mathbf{V} \cdot \begin{bmatrix} \mathbf{\Sigma}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \cdot \mathbf{U}^T = \sum_{i=1}^{W} \frac{1}{\sigma_i} \mathbf{v}_i \mathbf{u}_i^T. \quad (5)$$

In the computation of the pseudo-inverse of $\mathbf{S}$, the adverse effect of measurement noises may be aggravated due to the terms that contain reciprocals of small, near-zero singular values. The truncated pseudo-inverse method discards small singular values by truncating all the terms with index $\gamma \leq W$. That is, instead of using $\mathbf{\Sigma}$ in Eq. 5, we use $\mathbf{\Sigma}_\gamma = \mathrm{diag}(\sigma_1, \ldots, \sigma_\gamma)$, and the truncated pseudo-inverse of $\mathbf{S}$ can be written by $\mathbf{S}_\gamma^+ = \sum_{i=1}^{\gamma} \frac{1}{\sigma_i} \mathbf{v}_i \mathbf{u}_i^T$. To determine an adequate index γ, we use use the following criterion: the percentage accounted for by the first k singular values is defined by

$$\tau_k = \frac{\sum_{i=1}^{k} \sigma_i}{\sum_{i=1}^{W} \sigma_i}.$$

One may pre-determine a cut-off value, τ^* of cumulative percentage of singular values, and calculate γ to be the smallest integer such that $\tau_\gamma \geq \tau^*$. The threshold τ^* should be carefully chosen depending on the noise level of RSS measurements. We observe that the singular values $\sigma_{\gamma+1} \ldots \sigma_W$ of $\mathbf{S}$ are near-zero for $\tau^* = 0.98$. In our experiments, τ^* is set to 0.98. With the expression of $\mathbf{S}_\gamma^+$, SDM can be simply expressed as $\mathbf{T} = \mathcal{F}_d(\mathbf{D})\mathbf{S}_\gamma^+$, and can be on-line updated.

3.3 Implementation details

3.3.1 Preprocessing RSS samples

As anchor nodes continuously measure RSSs from other anchor nodes, we need to determine the value that represents the RSS at a given time instant, with which the SDM will be constructed and the locations of client nodes will be estimated. We use a median filter that takes the median value of RSS samples measured in a time interval of duration T_s, i.e.,

$$\bar{s}(t) = \mathrm{median}\{s(\tau) | t - T_s < \tau \leq t\},$$

where T_s is the time interval for the median filter and is set to 60 s in the experiments. As discussed in [47], use of a median filter can effectively deal with short-term variations in RSS measurements due to fast fading. In case of tracking the location of a client node, the maximum estimation delay is bounded by T_s.

3.3.2 Calibrating RSS at the same location

The only parameter we need to calibrate in the deployment phase is the RSS measured between two co-located

Table 1 RSS measurements between two co-located Linksys WRT54G wireless monitors with respect to different wireless network settings

Technology	Channel	Xmit power (mW)	RSS (dBm)
802.11b	2	28	−33.2651
802.11b	7	28	−33.4629
802.11b	2	56	−29.7642
802.11g	2	28	−34.3947

anchor nodes. Even though two anchor nodes are located at the same position, the RSS between themselves is not zero in practice (i.e., $s_{ii} \neq 0$ in Eq. 2 and is determined by factors such as the transmission power of wireless monitors. Note that this calibration is not performed for each wireless node to be localized, but only between anchor nodes of the localization system. Moreover, it is only performed *once* in the deployment phase. We measure the RSS between two co-located anchor nodes for 5 min and take the median value of the measurements as the *self-RSS*. The obtained values are listed in Table 1. If APs with different RF characteristics co-exist in the localization system (e.g., different brands of APs), the *self-RSS* for each pair of APs should be separatively measured.

3.3.3 Determining the self-distance

Recall that SDM uses the linearizing function to characterize the linear relationship between the RSS and the distance. If $\mathcal{F}_d = \log(d)$, its value for the distance from oneself (so-called self-distance, d_{ii}) is $-\infty$. If we drop the terms for the self-distance in Eq. 3, the problem becomes an optimization problem for an under-determined system. Instead, we use a small positive value ε to represent d_{ii}. Let $d_{\min}$ denote the smallest inter-anchor node distance, i.e., $d_{\min} = \min_{i,j\in\{1,\ldots,m\}}(d_{ij})$. To select ε, we linearize the logarithm function at $d_{\min}$ and obtain a linear function, whose y-axis intercept is set to be $\log(\varepsilon)$, i.e., $\varepsilon = d_{\min}/e$. In our implementation, we normalize the distance matrix $\mathbf{D}$ with a constant scale of $1/d_{\min}$ and replace its diagonal terms with $1/e$. On the other hand, if $\mathcal{F}_d = d$, the self-distance is simply set to be 0.

3.3.4 Determining the final location of the client

Once the distances from a client node to the anchor nodes are computed by SDM, a lateration algorithm is used to determine the location of the client node. We use the simple descent gradient method that minimizes the following objective function:

$$\mathcal{E} = \frac{1}{2}\sum_{i=1}^{m}(f_d(\tilde{\mathbf{x}}, \mathbf{x}_i) - \tilde{d}_i)^2,$$

where $\tilde{\mathbf{x}}$ and $\tilde{d}_i$ are the *estimated* location of the client node and the *estimated* distance to the ith anchor nodes computed by SDM, respectively. By differentiating the objective function, we obtain the following iterative equation for updating the estimate of $\tilde{\mathbf{x}}$:

$$\tilde{\mathbf{x}}[k+1] = \tilde{\mathbf{x}}[k] + \beta\sum_{i=1}^{m}\left(1 - \frac{\tilde{d}_i}{f_d(\tilde{\mathbf{x}}, \mathbf{x}_i)}\right)(\tilde{\mathbf{x}}[k] - \mathbf{x}_i),$$

where β is a constant step size and is set to 0.1. We set the initial estimate of $\tilde{\mathbf{x}}[0]$ to be the locations of the anchor node with the smallest distance estimate. Note that several other lateration algorithms are available in the literature [15, 29–32], and can be considered as an alternative to further improve localization accuracy.

3.4 Discussion

There are several deployment issues that are worthy of discussion:

3.4.1 Where to place wireless monitors

The APs in an 802.11 infrastructure are usually deployed to provide network connectivity, and their placement may not be optimal with respect to localization. Moreover, certain areas may not be covered by a sufficient number of APs. Under these cases, it is desirable to mount additional wireless monitors in advance and activate them as needed so that sufficient information can be collected to construct $\mathbf{S}$.

To determine where to place additional wireless monitors, we represent the initial topology of APs by an undirected simple graph, and find a set of locations $U = \{u_1, u_2, \ldots, u_k\}$ with the minimum cardinality, such that any position in the environment is covered at least by m APs or monitors (where $m = 4-5$ will suffice in practice). We leverage the work by Zhang and Hou [48] and use the necessary and sufficient k-coverage condition derived in [48] to select positions for placing additional wireless monitors.

3.4.2 Who should make the RSS measurement between a client and neighboring APs

Although in Sect. 3.2, we state that the task of measuring the RSSs between a client to be localized and its neighboring APs is performed by the client, it can also be performed by the APs/wireless monitors without any supports from client nodes. However, in this case, each AP has to continually measure the RSSs from nodes that are not

associated with itself, and this may cause the degradation of network throughput performance.

In our empirical study, we found that the measurement results in the two directions (RSS from a client to its neighboring AP, and vice versa) are often asymmetric. However, as the effects of physical wireless characteristics in the RSS measure have been *explicitly* figured in SDM, the asymmetry does not severely affect the localization accuracy, as long as all the measurements are made uniformly in one direction (from a client to its neighboring APs, or from the neighboring APs to a client).

3.4.3 How to deal with the co-existence issue of IEEE 802.11a/b/g interfaces?

Popular 802.11 WLAN infrastructure supports both 802.11b/g and 802.11a interfaces, with dual interfaces equipped at each access point (e.g., Cisco Aironet 1200). This, however, creates a deployment issue in the proposed localization work. 802.11a systems, defined in the 5 GHz UNII frequency band, suffer more in the case of line of sight (LOS) obstructions such as office walls, furniture, and human beings. In contrast, 802.11b/g, defined in the 2.4 GHz ISM frequency band, is comparatively less susceptible to the environmental effects. Therefore, RSS measurements and SDM construction under 802.11b/g may be "inconsistent" with those under 802.11a, and vice versa. In the current setting, we maintain two versions of localization mappings and differentiate clients based on their active interfaces. An interesting issue is how to correlate measurements made between these two types of interfaces, or at least reason about the discrepancy between the two types of measurements. This will be part of our future work.

4 Empirical evaluation

4.1 Software customization

To empirically evaluate the proposed localization system and compare it against PSS [15], we have completed a system prototype and deployed both approaches on the third floor in the west wing of Siebel Center for Computer Science. We have prototyped wireless monitors using open-source Linksys WRT54G Wireless-G routers. These Linksys routers are inexpensive (less than $60 per router), support IEEE 802.11b/g, and run Linux kernel 2.4 with its source codes [49] available under GPL. This allows us to customize the kernel sources and embed in it the Kismet 802.11 layer-2 wireless sniffer software [40], as well as other required monitoring functions including controlled channel scanning, neighborhood discovery, and selective capturing of management, control and data frames. The Ethernet interface at the Linksys router allows monitors to send their RSS measurements to the server and the server to control monitors. These wireless monitors take the role of APs for beaconing and making RSS measurements.

To customize the Linksys router software, we install and modify a third party Sveasoft Alchemy-pre7a firmware of Linksys WRT54G wireless routers to embed the (customized) Kismet wireless network sniffer software. Kismet software consists of two components, i.e., server and drone. Kismet drone runs on the WRT54G router, captures wireless packets, measures the received signal strength, and transmits through the Ethernet interfaces the captured packets (along with the measured signal strength) to a central host running the Kismet server. (In the current configuration, we have set up a single server that collects all RSS measurements from monitors.) The localization server connects to the Kismet server, obtains from the server the RSS measurements, implements the core functions of creating and maintaining the SDM, and displays the estimated locations of wireless clients.

4.2 Experimental setup

We have conducted experiments in three phases. In the first phase, we have deployed 8 wireless monitors in the west wing of Siebel Center as shown in Fig. 3 and

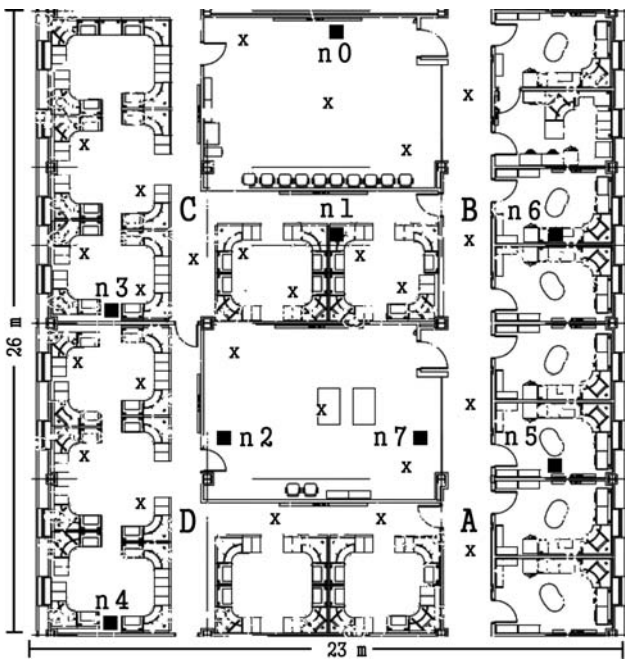


Fig. 3 Map of the west wing of Siebel Center for Computer Science consisting of 40 cubicles, 8 faculty offices, and 2 research labs. The black squares and the "x" marks denote the locations of the WRT54G wireless routers and the PDAs, respectively

measured the received signal strengths while varying the wireless network parameters such as IEEE 802.11b/g, the communication channels (2 (2.417 GHz) and 7 (2.442 GHz)), and the transmission power (28 and 56 mW) in order to evaluate the performance of the localization system in various environmental settings. Note that although 8 wireless monitors are deployed, localization is performed by using the RSS measurements at only 4–6 wireless routers (Sect. 3.4.1). In each experiment, one of the wireless monitors plays the role of a wireless client to be localized. In this manner, we can also easily investigate the effects of asymmetric wireless channels on localization accuracy.

In the second phase, we have used a PDA as the wireless device to be localized. The PDA runs Linux (Sharp Zaurus SL-5600) and is equipped with a wireless compact flash card (Linksys WCF-12). As the PDA client has different network characteristics from wireless monitors (WRT54G), e.g., the default transmission power of WCF-12 is 15 dBm (32 mW), we can evaluate the performance of the proposed localization system in a realistic environment with heterogeneous wireless devices.

In the third phase, a graduate student carries a PDA and walks along the path of [A → B → C → D → A] labeled in Fig. 3, and the location of the PDA is continuously monitored. Since the movement results in door opening and closing, we can evaluate the tracking performance of the proposed localization system with respect to robustness to environmental changes.

We have carried out experiments in different time intervals (morning, afternoon, evening, and late night) of the day in all three phases. We report below the results obtained from the experiments performed in the afternoon intervals when the environment is subject to the most human mobility and opening/closing of doors. The experiments are performed with either the proposed localization system or PSS [15] activated. The localization results for the proposed system are obtained with the use of the first-order linearizing function (i.e., $\mathcal{F}_d = d$) unless otherwise specified.

4.3 Experimental results

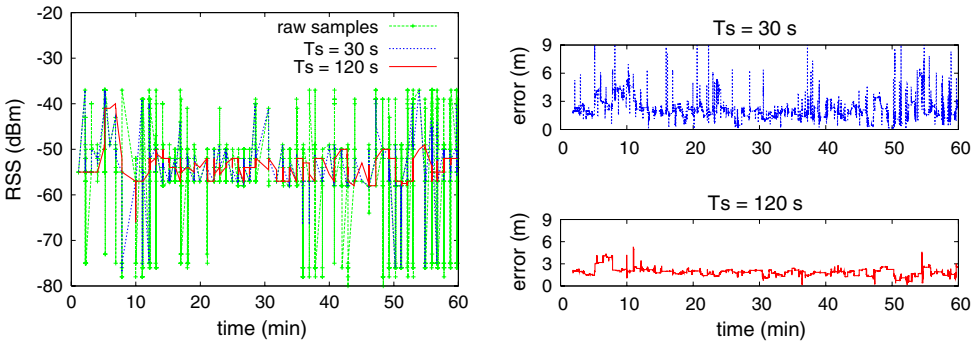
4.3.1 Impact of the time interval T_s for the median filter

In this set of experiments, we measure RSSs between *n1*, *n3*, *n4*, and *n7* in Fig. 3 for 1 h, construct the signal-distance mapping (SDM), and localize *n2* with the use of the RSS measurements made by *n2* itself. (That is, *n2* plays the role of a wireless device to be localized.) Figure 4(a) gives the strength of the signals sent by *n4* and measured at *n2*. As shown in the figure, the RSS measurements vary dramatically in the range of $-75 \sim -35$ dBm with a large number of signal breakdowns. However, the median filter mitigates the effect of RSS variation effectively. In the case of $T_s = 120$ s, the median filter takes more RSS samples and generates more stabilized estimate of RSSs. Figure 4(b) gives the localization results for *n2*. In both the cases of $T_s = 30$ and 120 s, the location estimates are quite accurate (with an localization error of 2 m on average). However, the estimate exhibits a large variance for $T_s = 30$ s.

To further investigate the effects of T_s on localization accuracy, we compute the mean and standard deviation values of localization errors for several sets of T_s and the probing rate at which beacon packets are send/received. As shown in Fig. 5, the mean and the standard deviation values become smaller as the probing rate and T_s increase. In fact, the number of samples taken for a fixed interval T_s depends on the probing rate. Note that the larger the value of T_s, the longer delay will be incurred in estimating the location of a client. In all the experiments henceforth reported, we set T_s to 60 s. Instead of using a fixed value of T_s, we may adaptively change its value according to the probing rate and the speed of mobile clients.

4.3.2 Localization accuracy

We evaluate the performance of the proposed localization system in the default wireless setting of the firmware. The WRT54G wireless routers are configured with the



(a) RSSs of *n4* measured at *n2*

(b) localization errors of *n2*

Fig. 4 Localization of a wireless client with the use of RSS measurements of four neighboring monitors. Location estimate of *n2* is generated at every 2 min with the RSS measurements made in the last T_s seconds

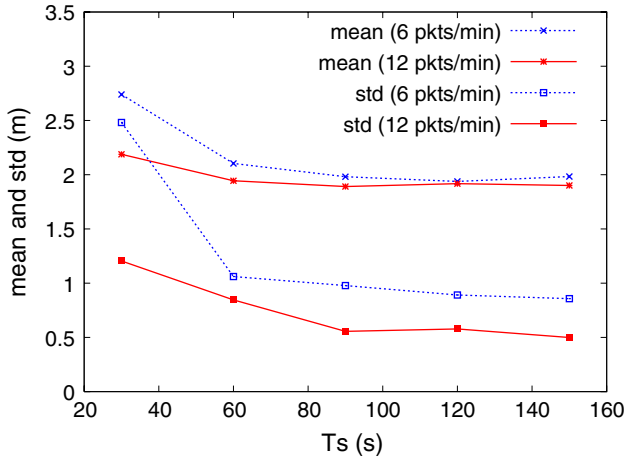


Fig. 5 The means and the standard deviations of localization errors versus T_s when the rates at which wireless monitors send beacon packets are 6 and 12 packets/min

Table 2 Localization errors in the default wireless network setting (in meters)

M	Algorithm	25%	Median	75%	Mean
(a) Localization errors of *n0–n7*					
6	SDM (client)	0.98	1.84	2.74	2.57
6	SDM (monitor)	1.47	2.31	2.86	2.92
6	PSS	2.82	6.15	10.44	6.92
5	SDM (client)	1.47	2.28	3.39	3.06
5	SDM (monitor)	1.93	2.63	3.61	3.22
5	PSS	2.45	5.28	8.78	6.82
(b) Localization errors of *n1*, *n2*, and *n7*					
6	SDM (client)	0.83	1.76	2.16	1.87
6	SDM (monitor)	1.47	2.16	2.52	2.05
6	PSS	3.74	6.80	7.41	7.09
5	SDM (client)	1.36	1.95	3.21	2.43
5	SDM (monitor)	1.76	2.56	3.37	2.71
5	PSS	3.68	5.97	8.14	6.40

following setting: 802.11b is used, channel = 2, and power = 28 mW. Figure 6 shows the cumulative percentage of localization errors for (a) *n0–n7* and for (b) {*n1*, *n2*, *n7*} when localization is carried out with the use of RSS measurements from 5 to 6 monitors (i.e., $m = 5$ and 6). For instance, in order to localize *n1*, we select m monitors except *n1*, compute the distances from *n1* to these m monitors, and determine the location of *n1*.

Several observation are in order. First, SDM gives better estimates when it uses RSS measurements from more wireless monitors. For example, as shown in Table 2(a), the median localization errors for the case of $m = 6$ are 1.84 and 2.31 m, and those for the case of $m = 5$ are 2.28 and 2.63 m. The accuracy obtained by SDM is better than that reported by rigorous off-line training-based algorithms (e.g., median error of 3 m and 97th percentile of 9 m in [50]).

Second, one predominant factor that affects the accuracy is the placement of these wireless monitors. The error incurred in localizing *n0* is actually much larger than that incurred in localizing, for example, *n1*, *n2*, and *n7*, because *n0* is not surrounded by wireless monitors (Fig. 3). As *n1*, *n2*, and *n7* are surrounded by several other monitors, they can be localized actually with the use of fewer monitors. As shown in Table 2(b) and Fig. 2(b), *n1*, *n2*, and *n7* can be more accurately localized. For example, the mean localization error incurred in localizing *n1*, *n2*, and *n7* with the use of 5 monitors is smaller than that incurred in localizing *n0–n7* with the use of 6 monitors. The issue of deploying and selecting wireless monitors has been addressed in [48, 51], and their algorithm will be incorporated in our next stage of deployment.

Third, the localization accuracy is also affected by which entities (a client or anchor nodes) make the RSS

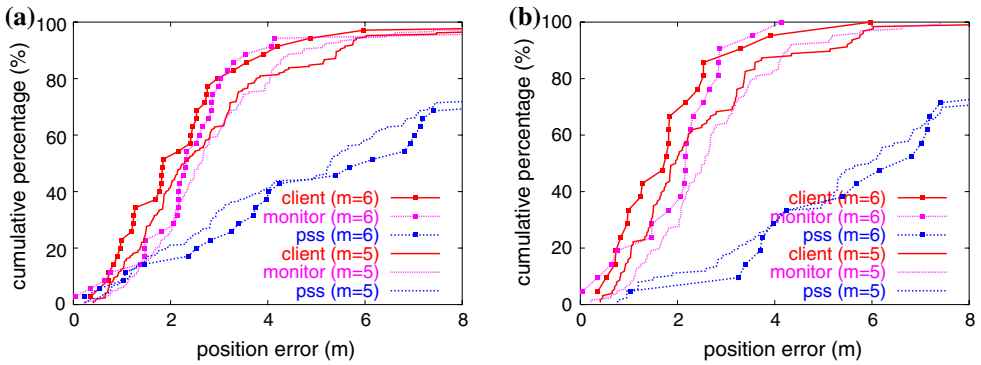


Fig. 6 Localization errors in the default wireless network setting. m is the number of monitors used for localization. Each wireless device is localized with the use of RSS measured either at the device itself (denoted by "client") or at the wireless monitors (denoted by "monitor"). The curve labeled with "pss" corresponds to the Proximity in Signal Space algorithm proposed in [15]. (**a**) Localization errors of *n0–n7*, (**b**) localization errors of *n1*, *n2*, and *n7*

measurement. As shown in Table 2 and Fig. 6, the localization accuracy achieved in the case of having the wireless client perform the RSS measurement is slightly better than that in the other case.

Last, as shown in Table 2 and Fig. 6, the localization error incurred in PSS is almost twice of that incurred in SDM in most cases.[3] This is, in part, due to the fact that PSS computes the distance from a wireless client to an AP with the use of RSS measurements between its closest AP and the destined AP, while SDM exploits measurements between all pairs of APs in the vicinity of the client. Also, the truncated SVD pseudo-inverse method introduced in Sect. 3.2 helps to mitigate the adverse effect of measurement noises.

4.3.3 Performance with respect to robustness

To evaluate the performance of the proposed localization system with respect to robustness against different network dynamics, we vary the wireless parameters of IEEE 802.11b/g, communication channels (2 (2.417 GHz) and 7 (2.442 GHz)), and transmission power (28 and 56 mW). To deal with these changes, the only operation we have to perform is to replace the RSS value for co-located wireless monitors with that listed in Table 1. For example, when the transmission power is set to be 56 mW, the RSS increases by about 4 dBm (theoretically 3 dBm). Figures 7–9 give the localization errors in the wireless network settings of (i) 802.11b, channel = 7, power = 28 mW; (ii) 802.11b, channel = 2, power = 56 mW; and (iii) 802.11g, channel = 2, power = 28 mW, respectively.

As expected, the experiment results do not differ significantly from those in the default network setting. This is attributed to the fact that construction of SDM does not rely on any assumption on physical wireless characteristics. An interesting finding is that the increase in the transmission power does not help to improve the localization accuracy, but rather degrades the performance. This is, in part, due to the fact that the increase in the transmission power makes the wireless channel more noisy (as a result of, for example, multi-path fading) and results in poor estimation of signal strengths.

4.3.4 Impact of the linearizing functions $\mathcal{F}_d$

We investigate the effects of linearizing functions on localization accuracy. Recall that in Sect. 3.2, we represent the relationship between the signal strength and the distance in a linear matrix form. Therefore, it is important to establish the linear relationship with the use of a proper linearizing function.

Figure 10 shows the comparison of estimation accuracy between $\mathcal{F}_d = \log(d)$ and $\mathcal{F}_d = d$ for $m = 6$. In Fig. 10(a), we observe that the localization errors for $\mathcal{F}_d = d$ are slightly smaller than those for $\mathcal{F}_d = \log(d)$, especially when the cumulative percentages are larger than approximately 40%. In contrast, in Fig. 10(b), the linearizing function of $\mathcal{F}_d = \log(d)$ gives better localization performance in almost the entire ranges. This implies that $\mathcal{F}_d = \log(d)$ is a better but more sensitive linearizing function (i.e., it gives accurate estimation but when its estimate is not accurate, the discrepancy is comparatively larger). This is due to the inverse function used in Eq. 4, which is an exponential function. This result implies that the selection of linearizing function does not severely affect the localization accuracy, which more relies on the capability of characterizing the anisotropic relationship between RSSs and geographic distances. Note that PSS characterizes the relationship between RSSs and distances with a linear function (like $\mathcal{F}_d = \log(d)$ in our algorithm), but gives a poor localization accuracy. However, for the sensitivity reason, we use the linearizing function of $\mathcal{F}_d = d$ in the experiments, which is simple to implement and achieves comparable estimation accuracy.

4.3.5 Performance in the case of device heterogeneity

A realistic wireless network consists of wireless devices of dramatically different characteristics, e.g., laptops, PDAs, VoWLAN phones, and Wi-Fi location tags. To evaluate the performance of the proposed localization system in a heterogeneous environment, we place PDAs in 25 different positions, measure at wireless monitors the strength of the signal emitted from these PDAs for 2 min, and localize PDAs. In this set of experiments, a "ping" program runs on the PDAs to transmit a wireless packet to its associated AP at every 1 s. Note that the rate at which the wireless traffic is generated is low enough to emulate typical network usage. Because the wireless channel is lossy, wireless monitors capture 73 wireless packets (out of 120) during the execution. Figure 11 gives the cumulative percentage of the errors incurred in localizing PDAs. The number of monitors that measure RSSs of the PDA varies from 4 to 7. As shown in Fig. 11, the median errors are at most 2.5 m under all cases. With a small number of RSS samples, the localization process converges and still achieves reasonable accuracy.

4.3.6 Performance of tracking mobile wireless devices

The ultimate goal of the proposed localization system is to accurately track moving wireless devices. If a mobile

[3] Gwon and Jain [15] has also proposed a lateration algorithm, called TIX, for computing the location of a client. We did not implement the TIX algorithm, but instead used the simple lateration algorithm given in Sect. 3.3.

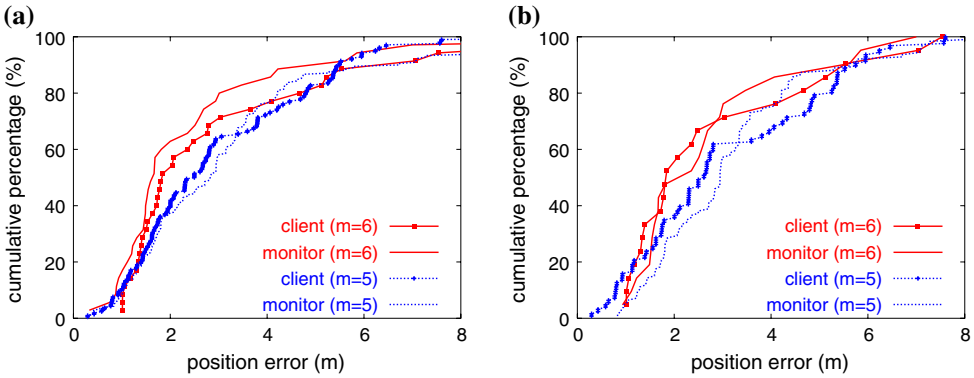


Fig. 7 Localization errors in the wireless network settings of 802.11b, channel = 7, power = 28 mW. (**a**) Localization errors of *n0*–*n7*, (**b**) localization errors of *n1*, *n2*, and *n7*

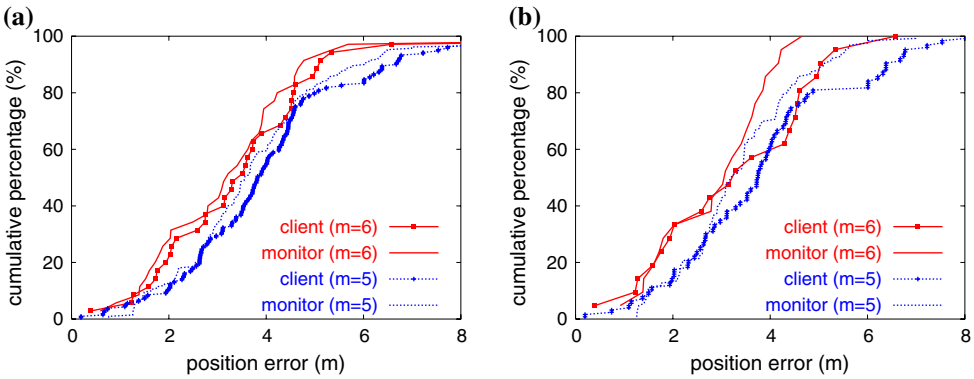


Fig. 8 Localization errors in the wireless network settings of 802.11b, channel = 2, power = 56 mW. (**a**) Localization errors of *n0*–*n7*, (**b**) localization errors of *n1*, *n2*, and *n7*

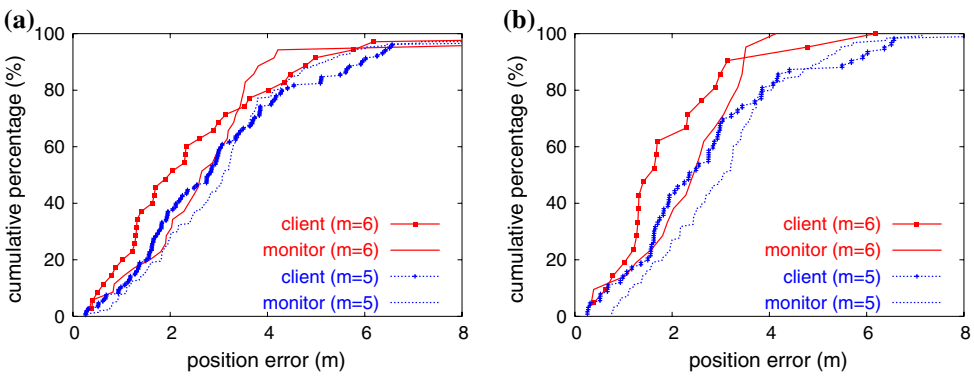


Fig. 9 Localization errors in the wireless network settings of 802.11g, channel = 2, power = 28 mW. (**a**) Localization errors of *n0*–*n7*, (**b**) localization errors of *n1*, *n2*, and *n7*

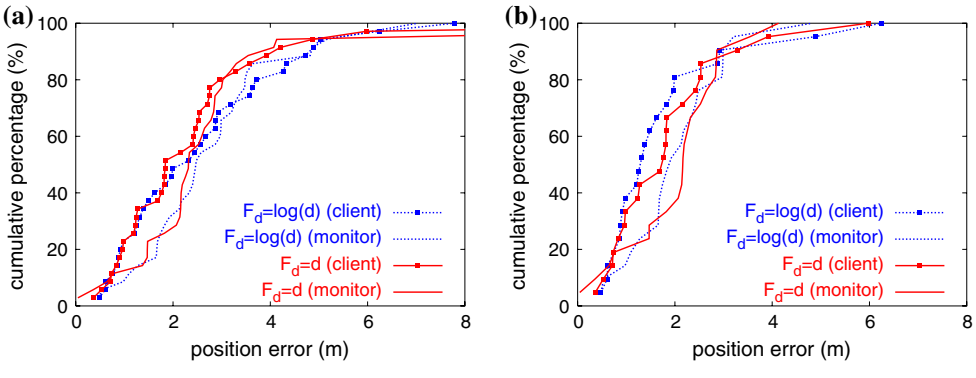


Fig. 10 Localization errors for two linearizing functions ($\mathcal{F}_d = \log(d)$ and $\mathcal{F}_d = d$) when $m = 6$ in the default wireless network setting. (**a**) Localization errors of *n0*–*n7*, (**b**) localization errors of *n1*, *n2*, and *n7*

device is still, wireless monitors in its vicinity can collect a number of RSS measurements (as the device transmits wireless packets) and constructs the SDM in due time. However, if a wireless device is mobile, the number of RSS measurements at intermediate locations may not be sufficiently large to render accurate location estimates. In this

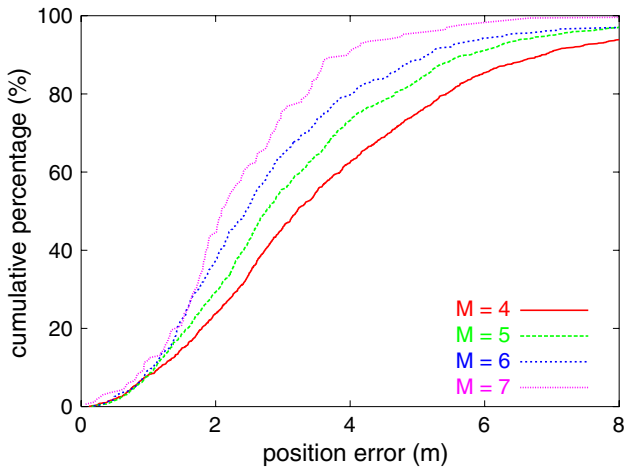


Fig. 11 Localization errors for PDAs placed at 25 different positions

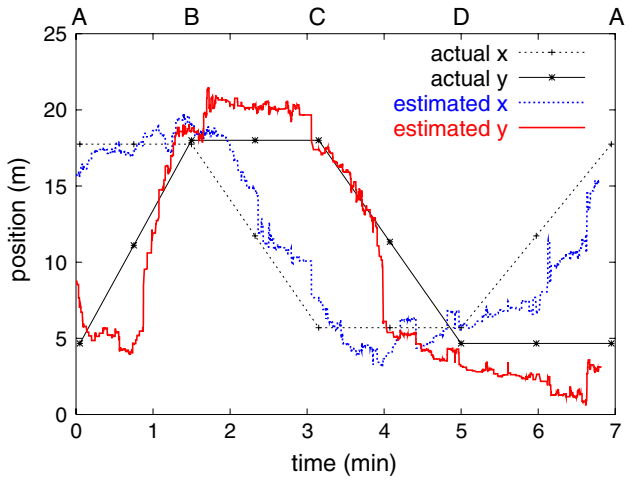


Fig. 12 The actual and estimated trajectory of a PDA (in the $[x,y]^T$-coordinate)

set of experiments, a graduate student carries a PDA and walks along the path of [A → B → C → D → A] labeled in Fig. 3. The PDA is in the process of downloading a file via an ftp application when the experiment is in progress. The location of the PDA is then continuously monitored with the use of 6 wireless monitors (*n0* and *n2*–*n6*). The locations of A and C are $[17.75, 4.67]^T$ and $[5.7, 18]^T$, respectively. Figure 12 gives both the actual and estimated trajectory of the PDA (in the $[x,y]^T$-coordinate) versus time. As shown in Fig. 12, the achieved accuracy for the PDA is within ±2.5–3 m. This suggests that as long as wireless devices are communicating with their associated APs in the process of moving,[4] this localization system can be used to track them with reasonable accuracy.

[4] This requirement can be relaxed if the wireless client is responsible for measuring RSSs of beacon messages from neighboring APs.

5 Conclusion

In this paper, we have established the theoretical base and developed a localization algorithm for building a *zero-configuration, robust, indoor localization and tracking* system to support location-based network services and management. Because of the widespread 802.11 deployment, we have leveraged the Wi-Fi infrastructure and addressed the major challenges of incorporating a fully-automated on-line calibration mechanism to characterize the relation between the received signal strengths (RSSs) and the geographical distances in the presence of environmental dynamics and measurement noises.

The localization algorithm takes as input the on-line measurements of RSSs (i) between 802.11 APs and (ii) between a client and its neighboring APs, and estimates the location of the client. The on-line RSS measurements among 802.11 APs are used to capture automatically and in real-time the effects of RF multi-path fading, temperature and humidity variations, opening and closing of doors, furniture relocation, and human mobility on the RSS measurements. The truncated SVD technique is then used to construct a mapping between the RSS measure and the actual geographical distance, with the objective of mitigating the adverse effect of measurement error, and yet retaining as much environmental information as possible. We have carried out an empirical study at Siebel Center for Computer Science, with the use of Linksys WRT54G wireless routers equipped with customized Kismet 802.11 layer-2 sniffer software. When clients perform the measurement of RSSs from APs, the accuracy we obtained is 2.28 m with 5 APs and 1.84 m with 6 APs in an office area of 600 m^2. The empirical results indicate that the proposed system is quite robust and responsive to environmental dynamics and gives accurate localization results.

An immediate research agenda is to investigate whether or not the localization error of 2.5 m can be further improved. There are several avenues along which we can pursue. First, as recognized in the empirical study, placement of wireless monitors has a significant impact on the accuracy. Although most of the APs in an operational building are fixed, the localization accuracy can be improved by deploying additional, inexpensive wireless monitors. Also, as mentioned in Sect. 4.3, the interval, T_s, (for which the median filter takes the median value of RSS measurements) represents a tradeoff between localization responsiveness and localization accuracy. Determining an adequate value for T_s, with consideration of the rate at which beacon packets[5] are gen-

[5] or in the case that the wireless monitors are responsible for measuring RSSs, the rate of which data packets are generated at a wireless client.

erated, would be an interesting problem. By a similar token, the rate at which beacon messages have to be generated in order to realize accurate *real-time* tracking is also a matter of future investigation.

Acknowledgement This work was supported by the basic research project through a grant provided by the Gwangju Institute of Science and Technology in 2007.

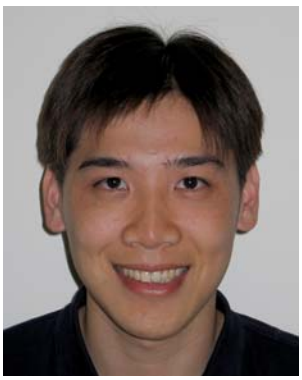
References

1. Lim, H., Kung, L.-C., Hou, J. C., & Luo, H. (2006). Zero-configuration, robust indoor localization: Theory and experimentation. In *Proceedings of IEEE INFOCOM*.
2. Goodstein, L. (2004). Location based service: Analysis of carrier spending, subscribers, devices and applications for handset-based and telematics services. Report form ABI research, http://www.abiresearch.com.
3. Nogee, A. (2001). Ready or not, mobile location technology is here! Cahners In-Stat Group.
4. Golmie, N. (2001). Interference in the 2.4 ghz band. In *Proceedings of International Conference on Applications and Services in Wireless Networks* (pp. 187–199).
5. Ekahau positioning engine. http://www.ekahau.com/products/positioningengine.
6. Castro, P., Chiu, P., Kremenek, T., & Muntz, R. (2001). A probabilistic room location service for wireless networked environments. In *Proceedings of Ubiquitous Computing*.
7. Haeberlen, A., Flannery, E., Ladd, A. M., Rudys, A., Wallach, D. S., & Kavraki, L. E. (2004). Practical robust localization over large-scale 802.11 wireless networks. In *Proceedings of ACM MOBICOM*.
8. Krishnan, P., Krishnakumar, A., Ju, W.-H., Mallows, C., & Ganu, S. (2004). A system for LEASE: Location estimation assisted by stationary emitters for indoor RF wireless networks. In *Proceedings of IEEE INFOCOM*.
9. Ray, S., Ungrangsi, R., Pellegrini, F. D., Trachtenberg, A., & Starobinski, D. (2003). Robust location detection in emergency sensor networks. In *Proceedings of IEEE INFOCOM*.
10. Pinpoint. http://www.pinpoint.com.
11. Ubisense. http://www.ubisense.net.
12. Priyantha, N. B., Chakraborty, A., & Balakrishnan, H. (2000). The cricket location-support system. In *Proceedings of ACM MOBICOM*.
13. Bahl, P., & Padmanabhan, V. N. (2000). RADAR: An in-building rf-based user location and tracking system. In *Proceedings of IEEE INFOCOM*.
14. Niculescu, D., & Nath, B. (2004). VOR base stations for indoor 802.11 positioning. In *Proceedings of ACM MOBICOM*.
15. Gwon, Y., & Jain, R. (2004). Error characteristics and calibration-free techniques for wireless lan-based location estimation. In *MobiWac '04: Proceedings of the Second International Workshop on Mobility Management & Wireless Access Protocols* (pp. 2–9). New York, NY, USA: ACM.
16. Want, R., Hopper, A., Falcao, V., & Gibbons, J. (1992). The active badge location system. *ACM Transactions on Information Systems, 10*(1), 91–102.
17. Harter, A., Hopper, A., Steggles, P., Ward, A., & Webster, P. (1999). The anatomy of a context-aware application. In *Proceedings of ACM MOBICOM*.
18. Priyantha, N. B., Miu, A. K., Balakrishnan, H., & Teller, S. (2001). The cricket compass for context-aware mobile applications. In *Proceedings of ACM MOBICOM*.
19. Li, T., Ekpenyong, A., & Huang, Y.-F. (2004). A location system using asynchronous distributed sensors. In *Proceedings of IEEE INFOCOM*.
20. Bahl, P., Padmanabhan, V. N., & Balachandran, A. (2000). Enhancements to the RADAR user location and tracking system. Microsoft Research, Tech. Rep. MSR-TR-2000-12.
21. Youssef, M., & Agrawala, A. (2003). Small-scale compensation for WLAN location determination systems. In *Proceedings of WCNC*.
22. Madigan, D., Elnahrawy, E., Martin, R. P., Ju, W.-H., Krishnan, P., & Krishnakumar, A. S. (2005) Bayesian indoor positioning systems. In *Proceedings of IEEE INFOCOM*.
23. Ladd, A. M., Bekris, K. E., Rudys, A., Marceau, G., Kavraki, L. E., & Wallach, D. S. (2002). Robotics-based location sensing using wireless ethernet. In *Proceedings of ACM MOBICOM*.
24. Tao, P., Rudys, A., Ladd, A. M., & Wallach, D. S. (2003). Wireless LAN location-sensing for security applications. In *Proceedings of ACM Workshop on Wireless Security*.
25. Ganu, S., Krishnakumar, A. S., & Krishnan, P. (2004). Infrastructure-based location estimation in WLAN networks. In *Proceedings of WCNC*.
26. Guvenc, I., Abdallah, C. T., Jordan, R., & Dedeoglu, O. (2003). Enhancements to RSS based indoor tracking systems using kalman filters. In *Proceedings of International Signal Processing Conference and Global Signal Processing Expo*.
27. Niculescu, D., & Nath, B. (2004). Error characteristics of ad hoc positioning systems. In *Proceedings of ACM MOBIHOC*.
28. Krishnakumar, A. S., & Krishnan, P. (2005). On the accuracy of signal strength-based location estimation techniques. In *Proceedings of IEEE INFOCOM*.
29. Niculescu, D., & Nath, B. (2001). Ad hoc positioning system. In *Proceedings of IEEE Globecom*.
30. Nagpal, R., Shrobe, H., & Bachrach, J. (2003). Organizing a global coordinate system from local information on an ad hoc sensor network. In *Proceedings of IPSN*.
31. Savvides, A., Han, C.-C., & Strivastava, M. B. (2001). Dynamic fine-grained localization in ad-hoc networks of sensors. In *Proceedings of ACM MOBICOM*.
32. Savvides, A., Park, H., & Srivastava, M. B. (2002). The bits and flops of the N-hop multilateration primitive for node localization problems. In *Proceedings of ACM WSNA*.
33. Savarese, C., Rabaey, J., & Langendoen, K. (2002). Robust positioning algorithms for distributed ad-hoc wireless sensor networks. In *Proceedings of USENIX Annual Technical Conference*.
34. Raykar, V. C., Kozintsev, I., & Lienhart, R. (2003). Position calibration of audio sensors and actuators in a distributed computing platform. In *Proceedings of ACM Multimedia*.
35. Shang, Y., Ruml, W., & Zhang, Y. (2003). Localization from mere connectivity. In *Proceedings of ACM MOBIHOC*.
36. Lim, H., & Hou, J. C. (2005). Localization for anisotropic sensor networks. In *Proceedings of IEEE INFOCOM*.
37. Bulusu, N., Heidemann, J., & Estrin, D. (2000). GPS-less low cost outdoor localization for very small devices. *IEEE Personal Communications, Special Issue on Smart Spaces and Environments, 7*(5), 28–34.
38. Doherty, L., Pister, K. S. J., & Ghaoui, L. E. (2001). Convex position estimation in wireless sensor networks. In *Proceedings of IEEE INFOCOM*.
39. He, T., Huang, C., Blum, B. M., Stankovic, J. A., & Abdelzaher, T. (2003). Range-free localization schemes for large scale sensor networks. In *Proceedings of ACM MOBICOM*.
40. Kismet. http://www.kismetwireless.ne.
41. Hjaltason, G. R. & Samet, H. (2003). Properties of embedding methods for similarity searching in metric spaces. *IEEE Transactions on Pattern Analysis and machine intelligence, 25*(5), 530–549.

42. Lim, H., Hou, J. C., & Choi, C.-H. (2005). Constructing Internet coordinate system based on delay measurement. *IEEE/ACM Transactions on Networking, 13*(3), 513–525.
43. Tang, L., & Crovella, M. (2003). Virtual landmarks for the Internet. In *Proceedings of ACM Internet Measurement Conference*.
44. Rappaport, T. S. (1996). *Wireless communications priciples & practice*. Prentice Hall, 1996.
45. Gencer, N. G., & Williamson, S. J. (1998). Differential characterization of neural sources with the bimodal truncated SVD pseudo-inverse for EEG and MEG measurements. *IEEE Transactions on Biomedical Engineering, 45*(7), 827–838.
46. Haykin, S. (1996). *Adaptive filter theory*. Prentice Hall.
47. Krishnakumar, A. S., & Krishnan, P. (2005). On the accuracy of signal strength-based location estimation techniques. In *Proceedings of IEEE INFOCOM*.
48. Zhang, H., & Hou, J. C. (2005). Maintaining sensing coverage and connectivity in large sensor networks. *Wireless Ad Hoc and Sensor Networks: An International Journal, 14*, 28.
49. Linksys WRT54G wireless-G broadband router souce codes. http://www.linksys.com/support/gpl.as.
50. Elnahrawy, E., Li, X., & Martin, R. (2004). The limits of localization using signal strength: A comparative study. In *Proceedings of IEEE SECON*.
51. Ray, S., Lai, W., & Paschalidis, I. (2005). Deployment optimization of sensornet-based stochastic location-detection systems. In *Proceedings of IEEE INFOCOM*.

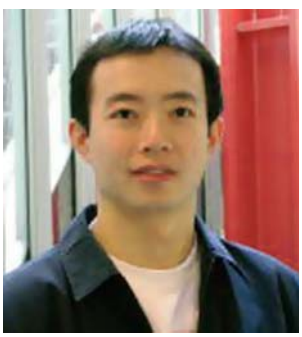
Author Biographies

Hyuk Lim received his B.S., M.S. and Ph.D. degrees from the School of Electrical Engineering and Computer Science, Seoul National University, Korea, in 1996, 1998, and 2003, respectively. He is an assistant professor in the Department of Information and Communications, Gwangju Institute of Science and Technology (GIST), Republic of Korea. From August 2003 to May 2006 he was a Postdoctoral Research Associate in the Department of Computer Science at the University of Illinois, Urbana. His research interests include network modeling and measurement, performance analysis and evaluation of sensor networks and mesh networks, and ubiquitous computing.



Lu-Chuan Kung received his B.S. degree from National Taiwan University in 2001. Currently, he is a Ph.D. precandidate in the Department of Computer Science, UIUC. His research interests include enabling techniques for large-scale network simulation, design and analysis of wireless protocols, and building wireless network applications.

Jennifer C. Hou was born on September 26, 1964 in Taipei, Taiwan. She received her B.S.E. degree in Electrical Engineering from National Taiwan University, Taiwan, ROC in 1987, M.S.E degrees in Electrical Engineering and Computer Science (EECS) and in Industrial and Operations Engineering (I&OE) from the University of Michigan, Ann Arbor, MI in 1989 and in 1991, and Ph.D. degree in EECS also from the University of Michigan, Ann Arbor, MI in 1993. She was an assistant professor in Electrical and Computer Engineering at the University of Wisconsin, Madison, WI from 1993-1996, and an assistant/associate professor in Electrical Engineering at Ohio State University, Columbus, OH from 1996–2001. She joined the University of Illinois Computer Science faculty in 2001. She was a principal researcher in networked systems and served as the director of the Illinois Network Design and Experimentation (INDEX) research group. She has supervised several federally and industry funded projects in the areas of network modeling and simulation, network measurement and diagnostics, and both the theoretical and protocol design aspects of wireless sensor networks. She has published (with her former advisor, students, and colleagues) over 160 papers in archived journals, book chapters, and peer-reviewed conferences. Her work on topology control and performance limits in wireless networks has been widely cited. She has been involved in organizing several international conferences sponsored by professional organizations such as ACM Mobicom, IEEE INFOCOM, IEEE MASS, and IEEE RTAS, as well as editor in archival journals and magazines such as IEEE Trans. on Computers, IEEE Trans. on Wireless Communications, IEEE Trans. on Mobile Computing, IEEE Trans. on Parallel and Distributed Systems, IEEE Wireless Communication Magazine, Elsevier Computer Networks, and ACM Trans. on Sensor Networks. She was a recipient of an ACM Recognition of Service Award in 2004 and 2007, a Cisco University Research Award from Cisco, Inc., 2002, a Lumley Research Award from Ohio State University in 2001, a NSF CAREER award from the Network and Communications Research Infrastructure, National Science Foundation in 1996–2000 and a Women in Science Initiative Award from The University of Wisconsin-Madison in 1993–1995. She was elected as an IEEE Fellow and an ACM Distinguished Scientist in 2007. Dr. Hou passed away on December 2, 2007 after a long and brave fight against cancer.



Haiyun Luo received the B.S. degree from University of Science and Technology of China, and the M.S. and Ph.D. degrees in Computer Science from University of California at Los Angeles. He is currently an assistant professor in Department of Computer Science, University of Illinois at Urbana-Champaign. His research interests include wireless and mobile networking and computing, embedded sensor networks, and network security. He is a member of the IEEE.